UNIVERSITY OF
LIVERPOOL

SIMULATION AND EVALUATION OF HV-CMOS PIXEL SENSORS FOR THE CLIC VERTEX DETECTOR

Thesis submitted in accordance with the requirements
of the University of Liverpool for the degree of
Doctor of Philosophy
by

Matthew Buckland*†

Supervisors:
Prof. Joost Vossebeld*
Dr. Dominik Dannheim†

*University of Liverpool, United Kingdom
†CERN, Switzerland

June 2018


FIAT

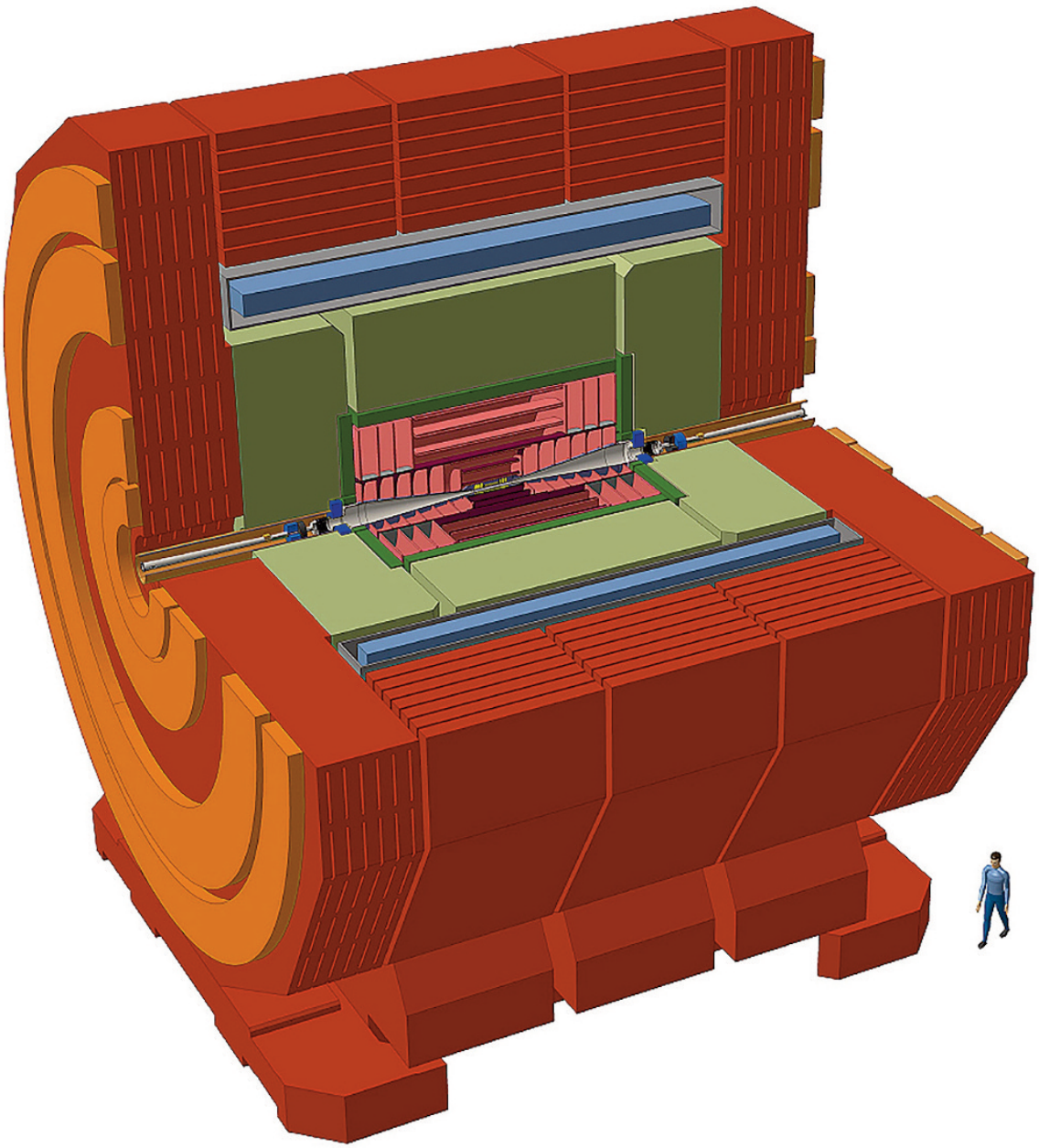
UNIVERSITY OF LIVERPOOL

Declaration

This thesis is the result of my own work, except where a specific reference to the work of others is made. This thesis has not been submitted for any other qualification to this or any other university.

Matthew Buckland

June 2018



An illustration of the CLICdet model with dimensions of 12:9min height and 11:4m in length. (see page 39)

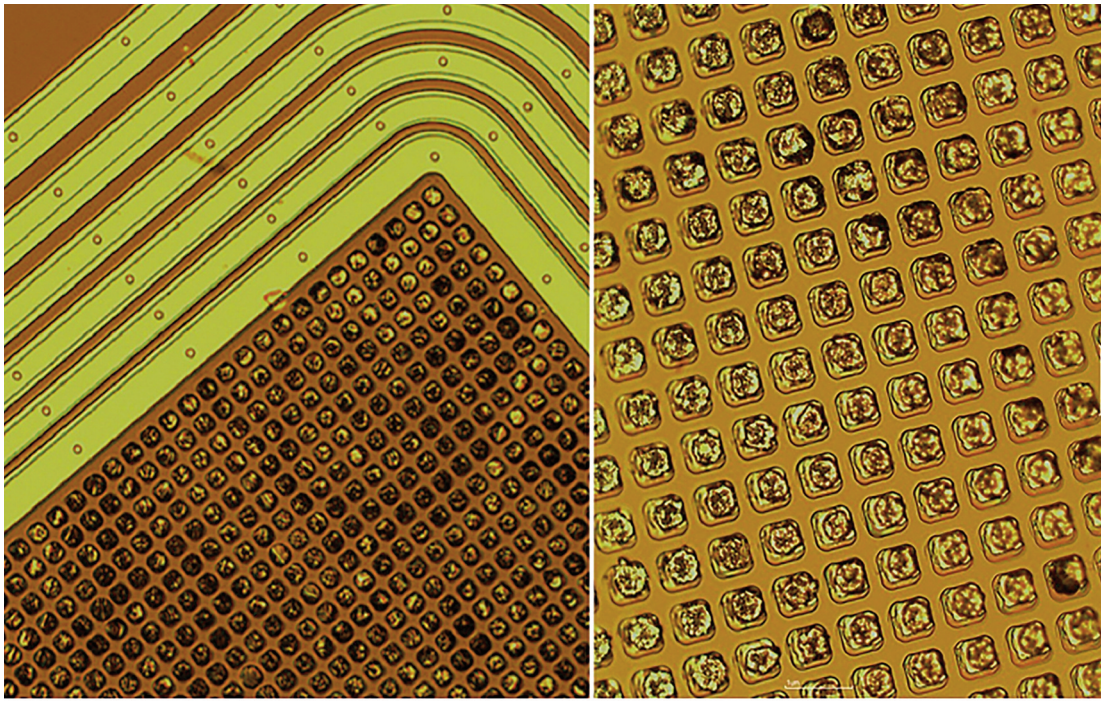
Abstract

The next steps in particle physics will involve colliders that are able to investigate the TeV energy scale so that many of the unanswered questions can be addressed. The Compact Linear Collider (CLIC) aims to do this through collisions of electrons and positrons at a high luminosity and at centre-of-mass energies of up to 3 TeV. In addition to the accelerator, a detector system is under development that targets precision physics measurements in an environment with a high rate of beam-induced backgrounds. One of the sub-detectors that faces particularly challenging requirements is the pixel vertex detector. To achieve its goals, hybrid readout chips either bump-bonded to planar sensors or capacitively coupled to High-Voltage CMOS (HV-CMOS) sensors, fabricated in a commercial 180 nm technology, are under study. Both of these sensor options have a small pitch of 25 x 25 μm^2 and are hybridised to 65 nm CLICpix readout ASICs. Initial investigations have shown the feasibility of such technologies, but further, and more detailed studies are needed. This is done through a series of simulations and measurements to assess the suitability of each technology for the CLIC vertex detector.

Simulations of the custom designed CCPDv3 HV-CMOS sensor have been carried out using the Sentaurus Technology Computer Aided Design (TCAD) simulation software. Firstly, a comparison of a 2D model and a resource intensive 3D model was carried out, showing an agreement within 15% of the sensor properties, such as electric field, depletion depth and charge collection. However, there is a difference of 40% for the capacitance, indicating 3D simulations are better suited to capacitance measurements. The 2D model was then expanded to include a multi-pixel model for studies of charge sharing.

Lab characterisations of planar sensor assemblies were undertaken to determine the quality of the bump-bonds and define regions with an acceptable level of working pixels for three assemblies. Calibrations were performed to convert the Time-over-Threshold (ToT) energy measurements and Digital-to-Analogue (DAC) threshold voltage steps into physical units. For the HV-CMOS assemblies, the analogue output of several pixels was compared to the ToT response of the CLICpix readout chip so that the simulation results could be converted to ToT and compared to the beam test data.

The performance of the HV-CMOS assemblies was assessed at the CERN SPS using 120 GeV/c secondary beams. This was done over an incident angle range of 0–80 ° and showed excellent efficiency above 99.7% and a spatial resolution of 5–7 µm after eta-correction was applied to correct for non-linear charge sharing. The measurements were then compared to the simulations, showing a good agreement for the current-voltage, breakdown and charge collection properties. The validated simulations were used to investigate possible prospects for improved performance that can be applied in future sensor designs. Two avenues of investigation were explored: increasing the bulk resistivity and biasing from the backside. The largest improvement was for a back bias model with a resistivity of 1 kΩ cm.



UBM and indium bumps deposited on (a) the readout chip (CLICpix) and (b) the slim-edge sensor prior to the flip-chip process. (see page 54)

Acknowledgements

To start, I would like to say that I am grateful to all of those whom I have had the pleasure of working with and to those that I have met throughout the duration of my PhD.

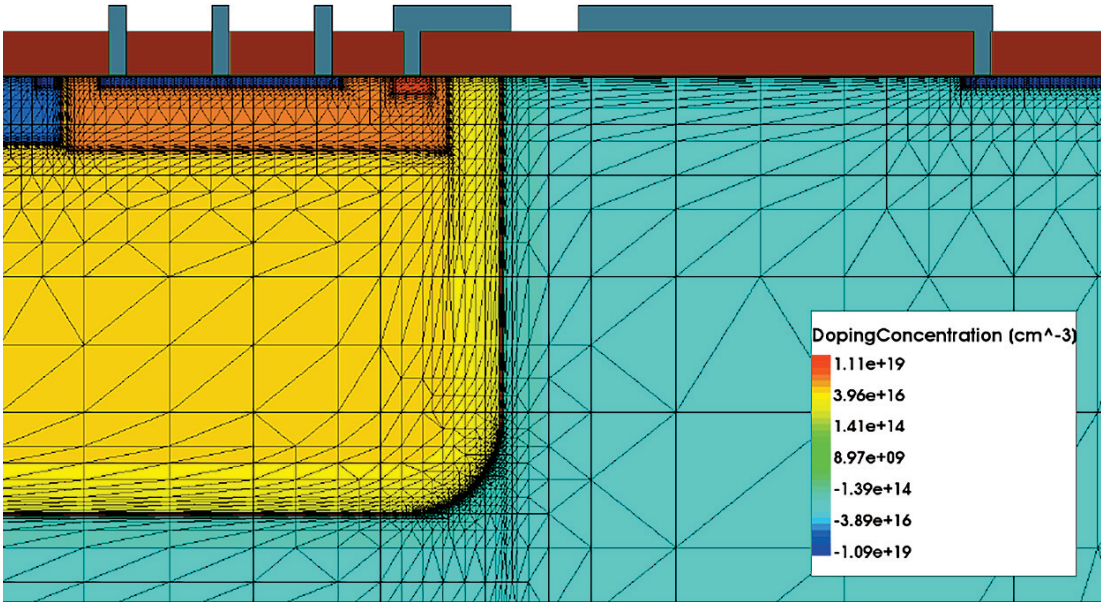
I would like to thank Joost Vossebeld for providing me with the opportunity to continue studying physics as a PhD student at the University of Liverpool and for being one of my supervisors. The other supervisor was Dominik Dannheim to whom I also owe thanks to. Their supervision has been vital to me in both supporting me during my PhD and developing myself as a physicist. They have also been sources of knowledge and encouragement which have driven me on to complete this body of work and I am very appreciative of all the effort that they put in to help me finish my PhD. Thanks should also go to those who provided me with invaluable support and guidance for specific areas of my work. For both my calibration and beam test work, Daniel Hynds has been very helpful and was always very generous to lend me his time in excellently explaining many of the concepts to me and providing advice through the many skype meetings we had. For my TCAD work, Mathieu Benoit first got me started on the simulation software and has provided the foundations upon which I was able to build upon, as well as setting up the AMS TCAD enthusiast group through which I benefited greatly. In addition to the above, I would also like to thank all the people of the CLICdp vertex community that I worked with for all their help along the way.

I also wish to express my gratitude to all of the people of the EP-LCD group who were very welcoming and made my time at CERN enjoyable. In particular, I would like to thank Lucie Linssen for allowing me the opportunity to work in the group and at CERN.

I am also thankful to all my fellow PhD students at the University of Liverpool and whom I met at CERN, for without their friendship, I am sure my time would not have been quite as fun. That can also be said of the CERN rugby team, who I shared many a great match with and who have allowed me to see that rugby union is all that bad.

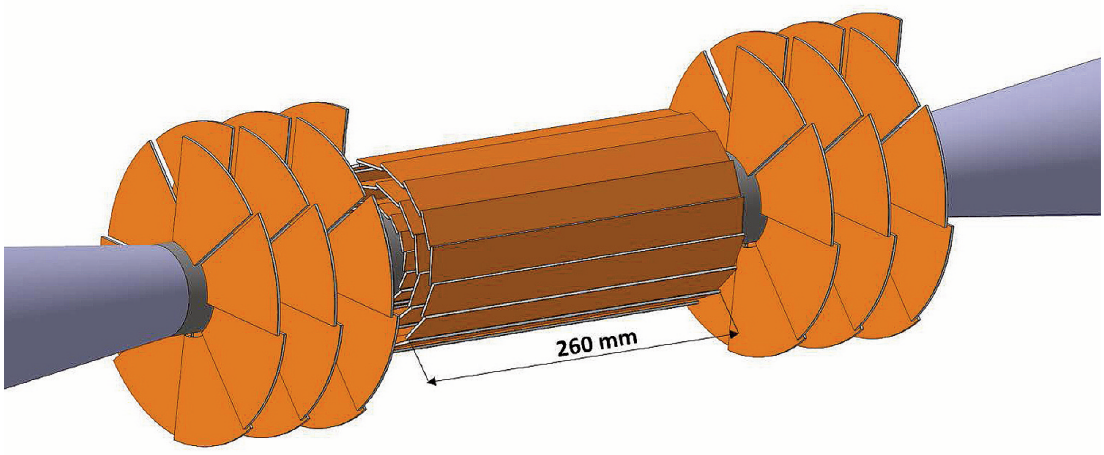
I would like to thank my parents and family for always providing me with the encouragement and support that I needed. My parents have also assisted me in the design of the layout of my thesis and provided crucial spelling and grammar checks.

Finally, I would like to say that without the support of all of the above, I would not have been able to complete this body of work and for that I am truly grateful.



An example of the mesh in silicon for a HV-CMOS pixel detector. The mesh is refined according to the gradient of the doping concentration and also refined at the boundaries between the silicon, aluminium and silicon-oxide. (see page 63)

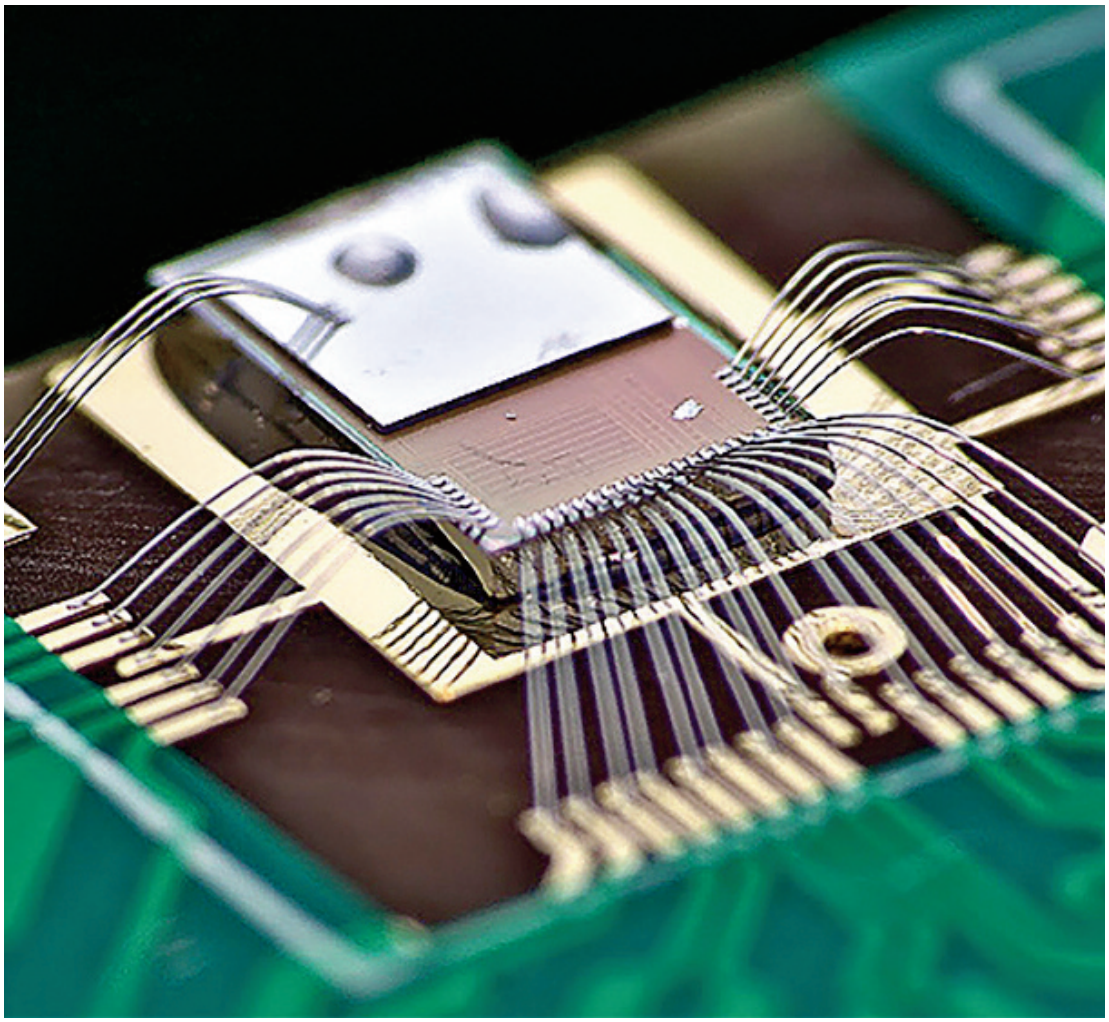
Contents



A schematic of the CLIC vertex detector showing the three double layers in the barrel and the spiral disks. (see page 51)

List of Figures

List of Tables



A close up of a planar sensor and the readout chip showing the wire bond. (see page 102)

CHAPTER 1
Introduction

In our current knowledge of physics there are four fundamental forces and they describe the interactions between all matter in the universe. Through many theories and experiments, the nature of this matter has been unearthed, revealing that everything is made up of a few fundamental particles. The Standard Model (SM) of particle physics provides a description of how the fundamental particles interact through three of the four fundamental forces (electromagnetic, strong and weak force). This theory has been robustly tested and has produced predictions such as the existence of the W and Z bosons, as well as the Higgs boson. These predictions have been observed in experimental results. However, this has also exposed some of the SM's weaknesses. It is not able to describe all the experimental phenomena at a subatomic level that have been observed. The challenge now is to fill in the knowledge gaps concerning the interplay between the four forces and all the elementary particles, to produce a complete theory of fundamental interactions.

Some of the phenomena the SM does not explain are: the baryon asymmetry, gravity (the fourth fundamental force), the existence of dark matter and dark energy. One of the more recent insights into the SM was the discovery of the Higgs boson at the Large Hadron Collider (LHC) [1, 2]. As well as providing evidence of the electroweak symmetry breaking mechanism [3, 4], it also raised the question of whether the new boson is the Higgs predicted by the SM, or a more exotic boson predicted by theories Beyond the SM (BSM). All these current challenges lead to a need for a comprehensive and precise exploration of the physics at the TeV energy scale, carried out by the LHC and future colliders.

One proposal for such a future collider is the Compact Linear Collider (CLIC) [5]. CLIC will be a high luminosity electron-positron linear collider operating at a centre-of-mass energy of up to 3 TeV and will be based at CERN. One consequence of a linear collider is that a very high acceleration gradient is needed, for CLIC this is $100\,\mathrm{MV\,m^{-1}}$. This will be achieved by using a novel two-beam acceleration technique. To maximise the physics reach of CLIC and to minimise power consumption and cost, it will be constructed in three stages, providing centre-of-mass energies of: 380 GeV, 1.5 TeV and 3 TeV. This provides a unique insight into important physics processes that occur in this wide range of energies and may lead to potential new discoveries. CLIC also aims to capitalise on the very precise measurements available at a lepton collider, building upon the knowledge acquired at the LHC.

To exploit the full physics potential of the CLIC machine, the detector must be able to cope with the experimental conditions such as a peak luminosity of $2 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and $\sqrt{s}$=3 TeV, which are considerably more challenging than the ones found at previous lepton colliders. This is due to the higher levels of beam-induced backgrounds and the rate and timing of the beams. To be able to deliver such a detector, technology beyond the current state-of-the-art is required along with comprehensive simulation studies and optimisation of the layout. All these are done to maximise the physics potential and minimise the negative effects due to backgrounds. The Conceptual Design Report (CDR) [6] showed that a wide range of high precision physics studies can be carried out with the CLIC detector following an intensive R&D programme.

One key component of the CLIC detector is the pixel vertex detector, located as close as possible to the interaction point in order to increase the measurement precision. It will provide precise and efficient tracking of charged particles and primarily measure the position of the decay vertices of heavy quarks and tau leptons. In order to achieve this, it must have low spatial resolution, full coverage, low mass, precise time-tagging and sufficient cooling. Currently no technology is able to simultaneously reach all these goals, so R&D is needed to produce a sensor and readout systems fit for such a purpose. The sensor concept under study is a hybrid readout chip coupled to a sensor. This is foreseen to be achieved in two ways, the readout chip is either bump bonded to planar sensors or capacitively coupled, via glue, to High-Voltage CMOS (HV-CMOS) sensors.

This thesis has contributed to the study of both sensor technologies. It has done so through extensive simulations of HV-CMOS sensor properties to better understand the phenomena that govern the functionality such as electric field, capacitance and leakage current. Beam test measurements were carried out to determine the suitability of the HV-CMOS sensor for the CLIC vertex detector. The simulations were validated with the measurements and subsequently used to predict possible prospects for improved performance for the next sensor generation. Furthermore, lab measurements are used to characterise and calibrate both the HV-CMOS and the planar sensors, to improve the understanding of the technologies.

The structure of this thesis is as follows. An introduction to silicon and its properties as well as how it is used as a charged particle detector is presented in Chapter 2. Next, an overview of CLIC is given in Chapter 3, detailing the physics motivation, the accelerator and the detector, followed by a separate chapter, 4, focusing on the vertex detector. After this, in Chapter 5, the set-up of the HV-CMOS sensor TCAD simulation is described along with the results of the various models used. Chapter 6 describes the method and results of the characterisation and calibration procedure of the two sensor technologies. Then, Chapter 7 provides a description of the beam test set-up and the method for analysing the data. After which, the results are shown and discussed. Finally, conclusions are given in Chapter 8.

CHAPTER 2
Silicon charged particle detectors

Silicon is the most commonly used semiconductor today due to its vast array of applications and its very well studied properties and fabrication processes. As a result, the majority of electronic devices are made using this material while other semiconductors such as GaAs are used for specific applications such as solar cells and lasers. In particle physics the detector material of choice is silicon for both the sensor and the readout chips.

This chapter will give a brief introduction to the theory aspects behind semiconductor properties such as: depletion, leakage current and ionisation, which have a large impact on the detector performance in areas such as spatial resolution and efficiency. It will also examine the equations that are of use for semiconductor device simulation, which will later be discussed in the context of simulations in chapter 5.1.2. An overview of the concept of silicon pixel detectors will also be introduced and explained along with a brief summary of the function of a readout chip.

2.1 Semiconductor properties

Knowledge of the properties of a semiconductor can be utilised when designing a particle detector as many effects can influence the performance. These include the doping levels, the generation of charge and the transport of charge within a semiconductor.

2.1.1 Crystal structures

The classification method for solid state materials is based on the arrangement of the atoms; this can occur in three ways: amorphous, polycrystalline and monocrystalline. The focus of this chapter will be on monocrystalline material as this is the most common structural form in solid-state devices fabricated for particle detectors. In a crystalline structure the atoms are arranged in a highly ordered three-dimensional array [7]. Crystal structures are important because they are very closely related to the intrinsic electrical properties of the material. A way to characterise the semiconductors is to look at their lattices, which are represented by their unit cell; the simplest repeating unit of a crystal [8]. The most common for semiconductors are the diamond lattice unit cell (Si) and the zincblende lattice unit cell (GaAs), shown in figure 2.1. In both cases there are eight atoms that all share their four valence electrons, forming four covalent bonds, and are characterised by the lattice constant a. The difference between the two is that

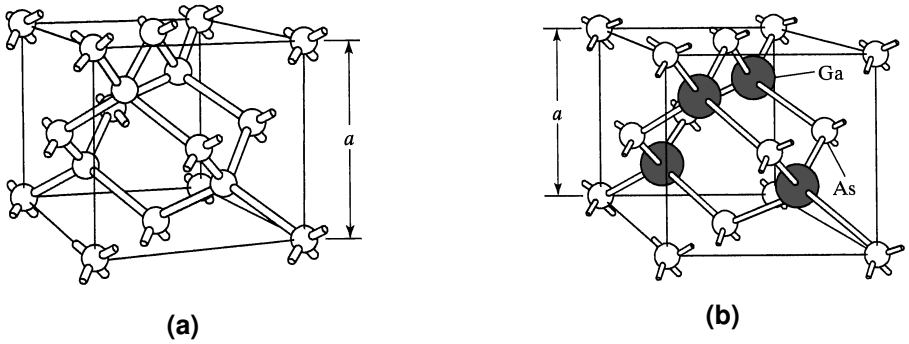


Figure 2.1: The most common semiconductor lattice unit cells (a) diamond and (b) zincblende [7]. The length a is the lattice constant.

in the diamond lattice all the atoms are identical, while the zincblende lattice consists of two atom types. An important property during fabrication is the identification of planes within the lattice. This is done by the use of Miller indices typically used to describe the surface orientation of the Si wafer. In semiconductor materials and in the simulations described in chapter 5, the crystal orientation is (111) a plane that cuts through the axes at x, y and z values of 1 [1]. The orientation is important to known during the fabrication process as the compatibility with processing equipment depends on accurate knowledge of the surface orientation.

2.1.2 Energy bands

One consequence of the crystal structure is that energy bands form, these consist of very closely spaced energy levels. This is due to the proximity of the atoms to each other within the material. If the atoms are sufficiently far apart, the electrons exist in discrete energy levels. As this distance is reduced, such as those found in a solid, these levels begin to merge, broaden and split. This is a result of the electron's wavefunctions overlapping with neighbours and, from the Pauli exclusion principle, the energy levels cannot be the same, producing closely spaced energy levels which form distinct energy bands. The band structure varies depending on the material and can be complex, so to simplify matters only two types of bands are considered: the conduction band and the valence band. The conduction band is the highest energy band containing electrons at temperatures above absolute zero. The valence band is the next lower band, separated by an energy gap from the conduction band. The electrons in the valence band are bound to the lattice sites while the electrons in the conduction band can travel through the lattice and contribute to the conduction of the material.

[1] The other typical orientation is (100).

The exact arrangement of these bands and whether they are empty, partially or completely filled determines the electrical conductivity of the solid. A completely empty band does not contribute to conduction as there are no electrons. For a partially filled band, there are only a few electrons and there are empty states with higher energies, this allows the electrons to gain energy, either thermally or electrically, and occupy these higher energy states, contributing to the conductivity. In the case of fully filled bands, there are plenty of electrons but no empty states so they cannot gain energy, hence do not contribute to the conductivity.

There are three types of classification for solids: conductor, semiconductor and insulators, whose band structures are shown in figure 2.2. In semiconductors the band gap is small enough (a few eV) that a few of the electrons in the valence band can reach the conduction band, via thermal excitations, yielding an almost full valence band and an almost empty conduction band. When an electron does this, it leaves behind an empty state called a hole. Both the electron and hole are involved in charge transport and current flow. The hole is an alternate description of the electron configuration in the valence band. It is used as a simplification because it is easier to describe a few empty states as opposed to keeping track of all the electrons. This means that the action of the electrons in a nearly filled valence band is modelled as particles with positive charge and positive effective mass occupying empty electronic states. The effective mass takes into account the effect of the periodic potential on the electron (hole) and it varies depending on the conduction (valence) band structure of the material.

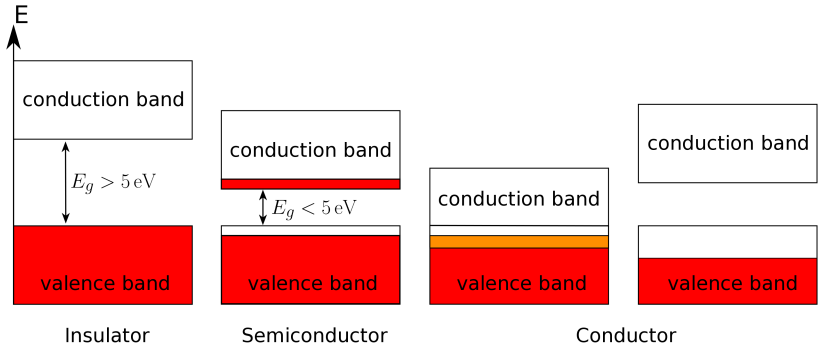


Figure 2.2: Energy band structure for insulators, semiconductors and conductors. The insulator has a completely full valence band and a completely empty conduction band, which are separated by a large band gap. For the semiconductor there is an almost full valence band and an almost empty conduction band, separated by a small band gap. There are two cases for the conductor: one where the two bands overlap and one where the valence band is half full.

2.1.3 Intrinsic semiconductor

There are three basic bond scenarios in a semiconductor, these are: intrinsic, n-type and p-type. The type of semiconductor is important to know because it determines quantities such as the number of carriers in each band and the energy distributions of the carriers, both of which are used for device modelling. The type of bond scenario can also be manipulated by the addition of impurities from different valence groups, more details will be given in the next section. In order to model these semiconductors, the starting point used is equilibrium; this occurs if the semiconductor has been left unperturbed for an extended period of time.

An intrinsic semiconductor is one in which the number of impurities is very low in comparison to the number of thermally generated free electrons and holes. As a result, the electron and hole carrier populations are equal. This is because for each electron generated by excitation, from the valence band to the conduction band, a hole will be created in the valence band. To calculate the concentration of free electrons, n, two elements need to be introduced [9].

The first element is the density of states, $g(E)$. This is the available number of states per unit energy per unit volume. This can be approximated by considering the density of states near the band edges, a region populated by carriers, for low enough temperatures and carrier densities. The electron in this region can be thought of as being confined to a three-dimensional potential well, with the well bottom being the minimum energy of an electron in the conduction band (E_C), and the well sides being formed by the edges of the crystal surfaces. Therefore, the density of states for electrons in the conduction band is given by:

$$g_c(E) = \frac{m_e^* \sqrt{2m_e^*(E - E_c)}}{\pi^2 \hbar^3}, \tag{2.1}$$

where m_e^* is the electron effective mass, E is the energy and $\hbar$ is the reduced Planck constant.

The second element is the Fermi function $f(E,T)$ that describes particles with half-integer spin, such as electrons. In thermal equilibrium it gives the probability that an available Fermion energy state will be occupied at a given energy E and temperature T. Since Fermions follow the Pauli exclusion principle, only one Fermion can fill each quantum state. As the electrons are added to an energy band, they occupy the states with the lowest energy first and stack up to build a Fermi sea of electron energy states.

The Fermi function is given by:

$$f(E,T) = \frac{1}{1+e^{(E-E_F)/kT}}\,, \tag{2.2}$$

where k is the Boltzmann constant and E_F is the Fermi level (electrochemical potential). The Fermi level, E_F is the energy at which the energy state has a 50% chance of occupation. This depends on the temperature, the mass of the particles and the density of particles. At $T = 0$ K the Fermi level is called the Fermi energy and defines the upper most filled energy state.

With the aforementioned elements determined, the free electron density can be calculated using:

$$n = \int_{E_c}^{E_{top}} g_c(E) f(E) \mathrm{d}E\,, \tag{2.3}$$

where E_{top} is the energy at the top of the conduction band, which can be approximated to ∞. The integration yields:

$$n = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} e^{(E_F-E_C)/kT} = N_C e^{(E_F-E_C)/kT}\,. \tag{2.4}$$

Analogously the density of free holes is given by:

$$p = 2\left(\frac{2\pi m_h^* kT}{h^2}\right)^{3/2} e^{(E_V-E_F)/kT} = N_V e^{(E_V-E_F)/kT}\,, \tag{2.5}$$

where m_h^* is the effective mass of holes and E_V is the maximum energy the electrons can have in the valence band. N_C and N_V are the effective densities of the conduction and valence band states respectively. Since the number of free electrons is equal to the number of holes it is often useful to define the intrinsic carrier concentration (n_i) by the law of mass action:

$$n_i^2 = np = N_C N_V e^{-E_g/2kT}\,. \tag{2.6}$$

This means that n_i is independent of the Fermi energy and only depends on the band gap energy, $E_g = E_C - E_V$, and the temperature.

In a semiconductor the Fermi function is the fraction of electrons that can bridge the energy gap into the conduction band and hence participate in electrical conductivity.

For an intrinsic semiconductor the number of electrons in the conduction band is the same as the number of holes in the valence band therefore, the Fermi energy is located in the centre of the band gap. By setting equations 2.4 and 2.5 equal to each other the Fermi energy can be calculated:

$$E_F = \frac{E_C + EV}{2} + \frac{kT}{2} \ln\left(\frac{N_V}{N_C}\right). \tag{2.7}$$

2.1.4 Doped semiconductor

Semiconductor devices typically do not use intrinsic material but instead the semiconductor is doped to obtain more desirable electrical properties. This process, which can be done either during crystal growth or later in specific areas, is done by introducing impurities into the crystal structure through the addition of foreign atoms. This has the effect of manipulating the carrier concentrations in one of two ways: either increasing the number of electrons to produce an n-type semiconductor or increasing the number of holes to produce a p-type semiconductor. In the n-type the electrons are said to be the majority carriers and the holes minority carriers, whereas in the p-type the holes are the majority carriers and the electrons are the minority carriers. The type of dopant depends on the foreign atom used.

For silicon (fourth valence group) this is done with elements from the third valence group (e.g. boron) or the fifth valence group (e.g. phosphorus). The third group atoms have one valence electron less than silicon and are called acceptors as they force a free hole to be created by accepting an electron from the valence band. While the fifth group atoms have one extra valence electron compared to silicon and are called donors as they produce a free electron. These two situations are illustrated in figure 2.3.

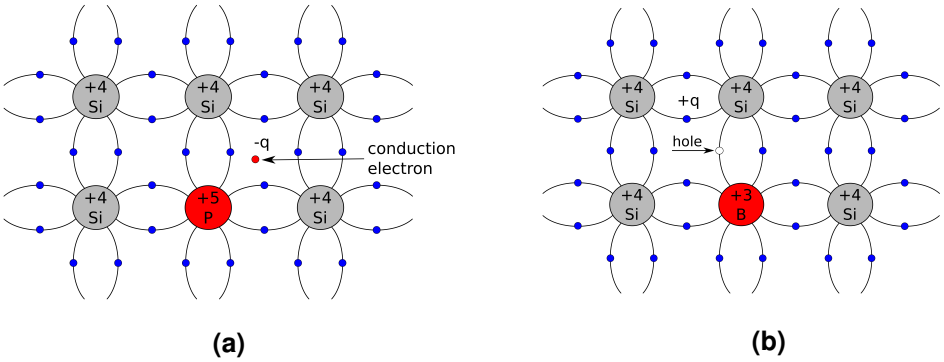


Figure 2.3: Bond representation of the crystal lattice structure for silicon with a (a) phosphorus atom, producing an n-type, and with a (b) boron atom producing, a p-type.

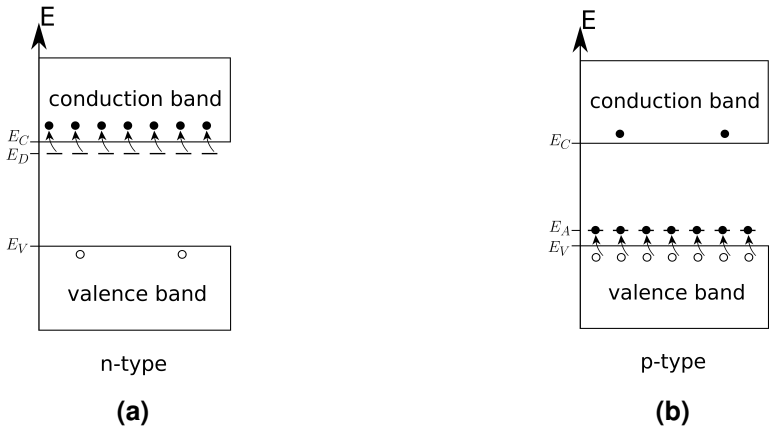


Figure 2.4: Energy band diagram showing the two doping scenarios for a semiconductor: n- and p-type. In the n-type the dopant atoms add an extra energy level (E_D) below the conduction band. In the p-type the dopant atoms add an extra energy level (E_A) above the valence band.

These dopants also have an effect on the energy bands for the semiconductor by adding additional energy levels in the band gap, shown in figure 2.4. For donors the energy level (E_D) is close to the conduction band edge, whereas for acceptors the energy level (E_A) is close to the valence band. At first these levels are filled by their respective particle, but at room temperature there is enough thermal energy for all the particles to bridge the gap to the next energy level, creating free carriers (ionisation).

2.1.5 The p-n junction

The p-n junction is an important structure in semiconductor physics as its properties can be exploited to create a wide variety of electrical devices including particle detectors. It consists of a semiconductor that has two oppositely doped regions next to each other, one side is n-type while the other side is p-type. Such a structure can be thought of as a diode as current mainly flows in only one direction.

Thermal equilibrium

Due to the concentration difference, the majority carriers on one side diffuse across the junction to the opposite side. Once there, they recombine and leave behind either ionised donors or acceptors near the junction. This means this region is free of charge carriers. The ionised donors and acceptors produce an electric field that opposes the diffusion current (more details about charge transport are given in chapter 2.1.7). This region is called the depletion region; any free carrier in this region will drift due to the electric field. Thermal equilibrium is reached when the two current flows balance each other out resulting in the Fermi levels of both semiconductors to line up, as illustrated

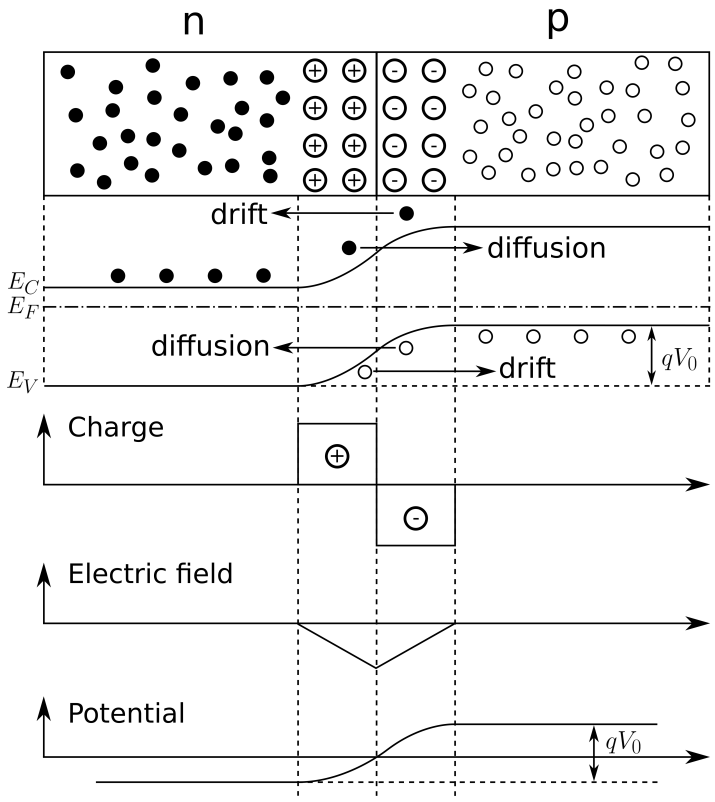


Figure 2.5: A schematic of a p-n junction in thermal equilibrium and its energy level configuration, including the charge, electric field and electric potential across the junction.

in figure 2.5. The electric field of the junction is characterised by the built-in potential:

$$V_0 = \frac{kT}{q} \ln\left(\frac{N_A N_D}{n_i^2}\right), \tag{2.8}$$

where q is the elementary charge and N_D and N_A are the concentrations of donors and acceptors, respectively. Figure 2.5 also illustrates the electric charge, electric field and electric potential across the p-n junction.

Reverse bias

The depletion region which occurs in a p-n junction is an attractive feature for particle detectors as the electric field found there can be used to collect the deposited charge. However, the depletion region is typically too small for such an application, around a few microns. To increase the size a change in the doping concentration of the semiconductors is needed or the application of a bias voltage across the junction. There are

two ways to do this, either apply a forward bias, where the positive terminal is connected to the p-side, or apply a reverse bias, where the negative terminal is connected to the p-side. Since the depletion region size decreases in the forward biasing case this scheme is not used for particle detectors and will not be further discussed. For the reverse bias case the majority carriers will be removed from either side and the depletion region will increase in size. The width, w, of the depletion region can be calculated by solving Poisson's equation. More details can be found in [7]:

$$w = \sqrt{\frac{2\varepsilon_r\varepsilon_0}{q}\left(\frac{1}{N_D}+\frac{1}{N_A}\right)(V_0 - V)}\,, \tag{2.9}$$

where ε_r is the relative permittivity of the substrate. The substrate is a thin slice of a semiconductor onto which other layers are built upon to make the device. ε_0 is the vacuum permittivity and V is the applied bias voltage.

Capacitance

When a p-n junction is formed it introduces a capacitance across the junction due to the stored charge at the depletion region. A variation in the depletion region will result in a variation in the capacitance. Under reverse bias conditions this capacitance is called depletion layer capacitance and is the dominant form of capacitance. It can be approximated as a parallel plate capacitor, with area A, and the depletion region acting as the separation between the plates:

$$C = A\frac{\mathrm{d}Q}{\mathrm{d}V} = \frac{A\varepsilon_r\varepsilon_0}{w} = A\sqrt{\frac{q\varepsilon_r\varepsilon_0}{2(V_0 - V)}\frac{N_D N_A}{N_D + N_A}} \approx A\sqrt{\frac{\varepsilon_r\varepsilon_0 qN}{2V}}\,. \tag{2.10}$$

The approximation comes from whether the semiconductor is doped with N representing either N_A (n-type) or N_D (n-type). A measurement of the capacitance can be used to extract the depletion depth, the doping concentration and the built-in potential. This is done by plotting $1/C^2$ as a function of applied reverse bias; the slope gives the doping concentration while the y-intercept gives the built-in potential. For each voltage the doping concentration and the depletion width can be determined giving a doping profile of the device. For a one-sided p-n junction, where one side is highly doped in comparison to the other, the slope is given by:

$$\frac{\mathrm{d}(1/C^2)}{\mathrm{d}V} = \frac{2A^2}{q\varepsilon_r\varepsilon_0 N}\,. \tag{2.11}$$

Hence the carrier concentration of the bulk is:

$$N = \frac{2A^2}{q\varepsilon_r\varepsilon_0}\frac{1}{\frac{\mathrm{d}(1/C^2)}{\mathrm{d}V}}\,. \tag{2.12}$$

2.1.6 Charge generation and recombination

When a semiconductor is perturbed from equilibrium the carrier concentration is modified, resulting in mechanisms that act to restore this, the so-called recombination-generation mechanisms. Due to the normal operation of semiconductor devices being at non-equilibrium conditions, recombination and generation play a key role in device operation.

The generation of free carriers can occur due to thermal processes, producing a leakage current. In addition to this there are also other methods for generation such as optical excitation and excitation by ionisation due to charged particles traversing the semiconductor. For this thesis optical generation will not be considered. All these generation processes liberate an electron from the valence band to the conduction band creating an electron-hole pair. The recombination processes involve an electron and hole recombining or annihilating each other. The interaction of particles in matter and the resulting carriers produced, are the basic principles behind a particle detector. Hence, it is useful to characterise the exact processes that occur.

Generation by charged particles

Charged particles passing through matter lose part of their kinetic energy by excitation of bound electrons and by ionisation. Since the particles are charged, the fundamental mechanism is the electromagnetic interaction; energy is mainly lost by the interaction of the charged particle with the electrons in the material. The energy loss per distance travelled (stopping power), $\langle \mathrm{d}E/\mathrm{d}x \rangle$, due to ionisation of a charged particle through matter is given by the Bethe-Bloch formula [10]:

$$-\left\langle \frac{\mathrm{d}E}{\mathrm{d}x} \right\rangle = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_ec^2\beta^2\gamma^2W_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right], \tag{2.13}$$

with:

K	$4\pi N_A r_e^2 m_e c^2 = 0.307075\,\mathrm{MeV\,cm^2}$
r_e	the classical electron radius
m_e	mass of the electron
c	speed of light

z	charge number of incident particles
Z	atomic number of absorber
A	atomic mass of absorber
β	velocity of the traversing particle in units of the speed of light
γ	Lorentz factor
W_{max}	the kinetic maximum energy which can be transferred to an electron by a charged particle
I	mean excitation energy
δ	density-effect correction.

This is an average and takes into account different effects but is not the dominant process for electrons due to their small mass (amongst other properties). In this case equation 2.13 is modified. In most practical cases the formula is dependent on the velocity β with only a minor dependence on the mass introduced from W_{max}.

The shape of the curve produced by the formula is illustrated in figure 2.6. It shows the mass stopping power, which is the stopping power per unit density. Details of all the regions and contributions to this curve can be found in [10]. Particles whose energy loss is found close to the minimum of the Bethe-Bloch formula are said to be minimum ionising particles (MIP). In semiconductors the number of electron-hole pairs created by a charged particle passing through the material is dependent on the average energy needed to create a pair. For silicon this is 3.6 eV and results in around 80 e-h pairs per micron being produced.

Whilst equation 2.13 is useful for certain applications, in practice the average energy loss given is not a very accurate quantity when dealing with silicon detectors. A more relevant quantity for silicon detectors to use is the most probable value (MPV) of the energy loss given by Landau theory [11]. This theory describes the fluctuations in energy loss. Another important quantity is the shape of the probability density function. A graph showing the Landau function as a function of electronic energy loss per unit length is shown in figure 2.7. The width of the distribution depends on the thickness of the material, with the fluctuations increasing the thinner the sensor is. This is especially true towards the tail of the energy distribution where high energy losses take place and the average becomes skewed due to rare events with large single-collision energy deposits, elongating the tail. One example are the so-called δ-electrons, where secondary electrons are created from the initial charged particle and have enough energy to produce further ionisation.

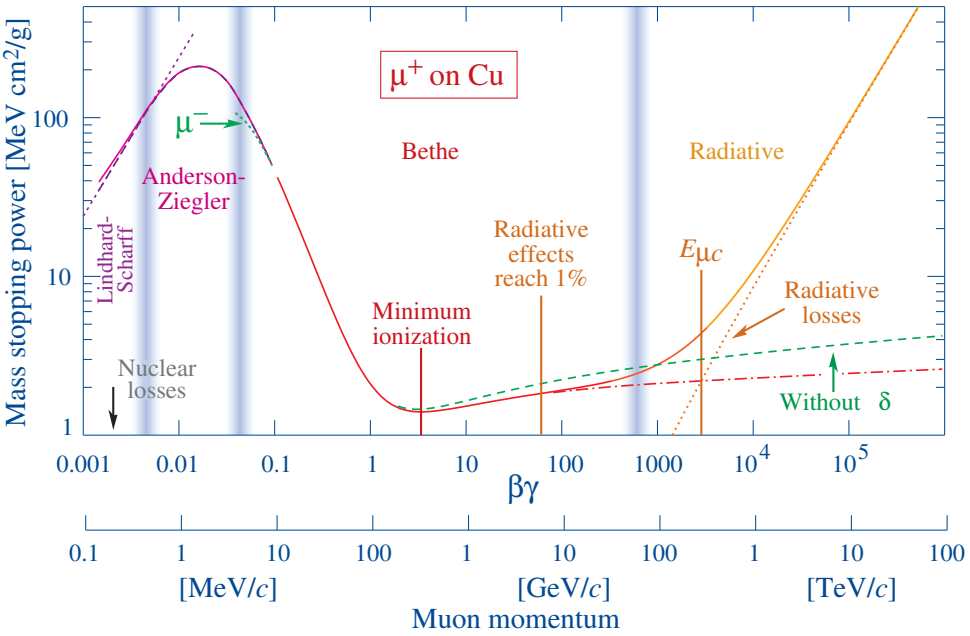


Figure 2.6: The mass stopping power of muons in copper as a function of $\beta\gamma$ [10]. The solid curve represents the energy loss with all effects added, while the dashed and dotted ones represent the individual contributions.

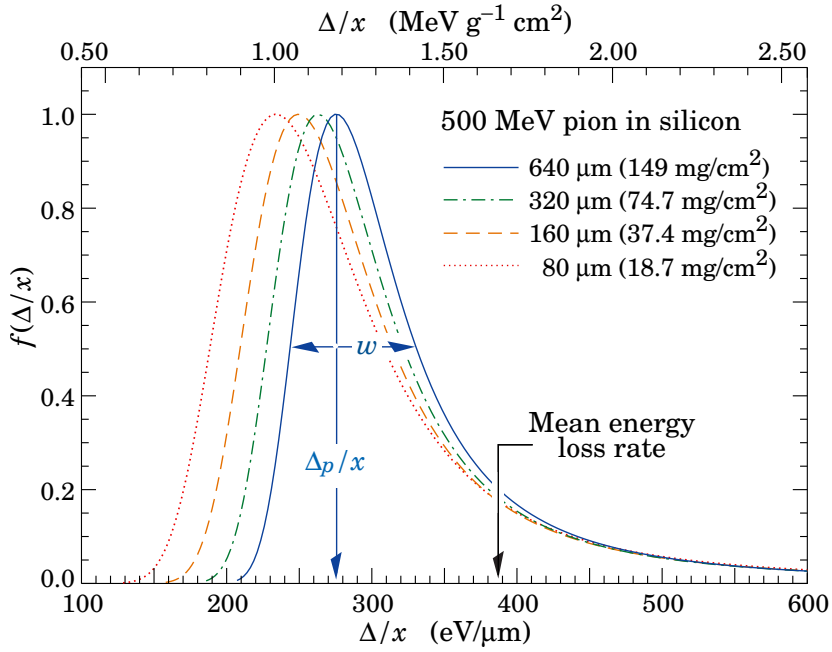


Figure 2.7: The Landau probability density function in silicon for 500 MeV pions, normalised at the most probable value (MPV), indicating the difference between the mean and the MPV [10]. w is the full width at half maximum.

The shape of the charge deposition and the penetration of the particle in the material are important quantities, that also depend on the type of radiation. High energy ($\sim$100 MeV) charged particles have enough energy to traverse the material at a near constant (relativistic) velocity producing uniform ionisation along their path. The density of electron-hole pairs created is almost constant with energy but proportional to the square of the charge of the particle. Another type of radiation is β radiation, which consists of a single electron. It penetrates deeply or possibly passes straight through due its low mass and can be considered as a MIP. For low energy particles such as alpha particles the energy loss due to ionisation is large, resulting in a very small penetration depth. As such particles pass through the material, they ionise more as their velocity decreases, with a peak at the stopping point.

Another important effect when considering particles passing through material is scattering off the nuclei by the Coulomb interaction. The particle is affected by many small angle scatters that produce a net scattering described by a Gaussian distribution with an RMS of [10]:

$$\theta_0 = \frac{13.6\,\mathrm{MeV}}{\beta cp} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \ln\left(\frac{x}{X_0} \right) \right], \tag{2.14}$$

where p is the momentum and x/X_0 is the thickness of the material in radiation lengths. The radiation length X_0 is the mean length over which a high energy electron loses $1/e$ of its initial energy in a specific material.

Lifetime of carriers

The characteristic quantity used to define the transient behaviour from a perturbed state back to equilibrium is called the charge carrier lifetime. This quantity can depend on numerous quantities such as material quality, doping level and the injection of carriers. There are two types of charge carrier lifetimes: the recombination lifetime τ_r, the average time taken before a minority carrier recombines (typical values are $\sim 1 \times 10^{-6}$ s for detector grade silicon and used in simulations); and the generation lifetime τ_g, the average time to generate an electron-hole pair.

Recombination and generation

Recombination and generation processes are terms used to describe the change in the carrier concentrations as a function of time. The type of process that occurs depends on the relationship of np to n_i. If there is an excess of carriers ($np > n_i^2$) caused by e.g. a pulse of light, then the dominant process is recombination as this reduces the number of carriers and will eventually return the system back to equilibrium. This is

done by the recombination of the excess of minority carriers with the majority carriers. The time it takes for this to happen is exponential and depends on the characteristic time τ_r. Recombination processes can be classified as either radiative or non-radiative. When the two carriers recombine, the difference in energy between the initial and final state is released. The energy is either released by a photon, radiative, or it is given to another electron or emitted as phonons, non-radiative.

In the case of a lack of carriers ($np < n_i^2$) the rate of the recombination processes becomes very small, allowing generation to dominate, resulting in an exponential return to equilibrium defined by τ_g. This is achieved by producing new carriers by exciting electrons from the valence band to the conduction band.

Band-to-band recombination and generation can be heavily suppressed if the material has an indirect band gap, such as Si, compared to a direct band gap, such as GaAs. For an indirect band gap the minimum of the conduction band does not have the same crystal momentum (wave vector $\mathbf{k}$)[1] as the maximum of the valence band, this can be thought of as the two not lining up in k-space. Therefore, a change in momentum is also required by an indirect band gap, which can be provided by a phonon (more details can be found in [9]). Alternatively, the recombination and generation for an indirect band gap material can occur via intermediate states that exist in the band gap, introduced through defects in the crystal and impurities in the semiconductor. These intermediate states become especially important after irradiation and can lead to an increase in the recombination, producing a loss in collected charge and an increase in generation, producing a larger leakage current [12].

2.1.7 Charge transport

Once charge has been generated, the motion of the free carriers via different charge transport mechanisms leads to a current. The motion due to an electric field, produced by an externally applied voltage, is called drift. Another method, called diffusion, is the movement due to a concentration gradient, moving from high carrier density to low carrier density, and is associated with the random thermal motion of the carriers. In both cases, the same free particles are involved and the same scattering mechanisms occur.

Drift

Drift is the motion of a charged carrier under the influence of an electric field. This accelerates the carriers and is often interrupted due to scattering events caused by

[1] The crystal momentum is a momentum-related constant describing the motion of the carriers in the crystal lattice and incorporates the electron's interactions in the crystal, $\mathbf{p}_{crystal} = \hbar\mathbf{k}$ [9].

defects in the crystal, thermally agitated lattice atoms or doped atoms. As a result, carriers do not move in straight paths along the electric field lines. Instead, their path changes direction many times on a microscopic scale, while their overall movement is in accordance with the electric field. Averaging over one type of carrier, the motion is described by the drift velocity:

$$\upsilon_n = -\frac{q\tau_c}{m_e^*}\mathscr{E} = -\mu_n\mathscr{E} \qquad \text{for electrons}, \tag{2.15}$$

$$\upsilon_p = \frac{q\tau_c}{m_h^*}\mathscr{E} = \mu_p\mathscr{E} \qquad \text{for holes}, \tag{2.16}$$

where $\mathscr{E}$ is the electric field, τ_c is the average time before scattering (mean free time) and μ is the mobility of the carrier. For small electric field values, the drift velocity is proportional to the electric field, figure 2.8. When the electric field becomes larger, the linear relationship breaks down and as the velocity increases so do the number of collisions leading to a saturation in the velocity, v_{sat}. In silicon at room temperature, the saturation velocity for electrons is $\sim 1 \times 10^7\,\mathrm{cm\,s^{-1}}$.

The mobility is used to characterise the carrier transport due to drift. It can be thought of as a measure of the ease of carrier motion when pulled by an electric field. It only depends on the mean free time. Therefore, the more scattering that takes place the lower the mobility. Due to this the mobility depends on temperature, electric field, doping concentration and the concentration of lattice imperfections. Electrons have about three times the mobility of holes. The mobility is also important for the conductivity σ and resistivity ρ of a semiconductor. Resistivity is a measure of how strongly the material opposes the flow of an electric current. It is an important parameter when designing

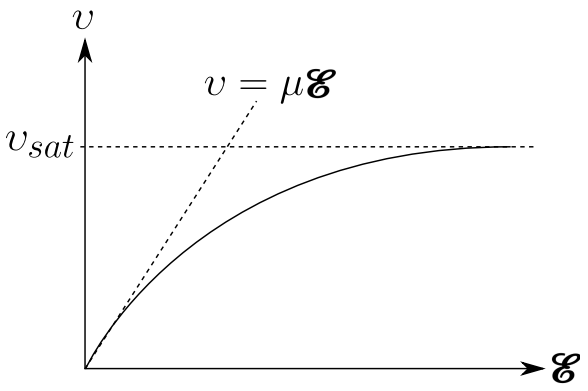


Figure 2.8: The velocity of a carrier in a semiconductor as a function of electric field.

a semiconductor particle detector as it determines the size of the depletion region, and in turn increasing the charge collection. The current produced by drift relates all these quantities:

$$\mathbf{J}_{\text{drift}} = \sigma\mathscr{E} = \frac{1}{\rho}\mathscr{E} = q(\mu_n n + \mu_p p)\mathscr{E}\,. \tag{2.17}$$

Therefore, the resistivity is given by:

$$\rho = \frac{1}{q(\mu_n n + \mu_p p)} \approx \frac{1}{qN\mu} \text{ (for doped Si)}\,. \tag{2.18}$$

Typical values of the resistivity for detector grade silicon are of the order of $\sim$1 kΩ cm.

Diffusion

In the absence of an electric field, and at temperatures greater than 0 K, the carriers spread out due to their random thermal motion. Overall, this results in no net movement. However, if there is a concentration gradient, then the carriers will move from the regions of high carrier densities to the regions of low carrier densities, eventually resulting in a uniform distribution. Since the carriers are charged, their overall movement produces a current, the diffusion current:

$$\mathbf{J}_{\text{n,diff}} = -D_n\nabla n = -\frac{kT}{q}\mu_n\nabla n \qquad \text{for electrons}\,, \tag{2.19}$$

$$\mathbf{J}_{\text{p,diff}} = \quad D_p\nabla p = \quad \frac{kT}{q}\mu_p\nabla p \qquad \text{for holes}\,. \tag{2.20}$$

$D_{n/p}$ are the diffusion constants for electrons and hole respectively and ∇n and ∇p are the gradients of the electron and hole concentrations.

2.1.8 Current and breakdown

In the idealised case the Shockley equation [13] describes the current in the p-n junction. One major assumption is that there is no charge generation in the depletion region. Due to particle detectors being operated in reverse bias, the Shockley equation is not sufficient to describe the current in this mode of operation. Therefore, the focus of this section will be the current and breakdown in the reverse bias scheme. Under reverse bias equilibrium is no longer achieved, hence there will be leakage current present. One component of this is the diffusion of carriers from the undepleted region to the depletion zone. However, the main component of this is the generation of carriers through

thermal excitation. This is because the reduction in the carrier concentration for reverse bias ($np \ll n_i^2$) leads to generation processes dominating in the depleted volume. The total reverse current density J_R is therefore given by the thermal generation rate of carriers G_{th} multiplied by the depletion width:

$$J_R = qG_{th}w = q\frac{n_i}{\tau_g}w \approx q\frac{n_i}{\tau_g}\sqrt{2\varepsilon_r\varepsilon_0\rho\mu V} \tag{2.21}$$

where τ_g is the carrier generation lifetime. Since the dominant part of the current is the thermal generation, it is important to characterise the temperature dependence to minimise the undesirable leakage current. This relationship is given by [13]:

$$J_R \propto T^2 e^{-E_g(T)/2kT}. \tag{2.22}$$

If a high enough reverse bias is applied, the electric field becomes large enough to cause breakdown. This is when there is a rapid increase in current at a certain voltage, called the breakdown voltage, as shown in figure 2.9. There are three mechanisms that cause breakdown: thermal instability, Zener breakdown and avalanche multiplication.

Thermal instability occurs because of the heat dissipation caused by the reverse current, producing an increase in the device temperature. Consequently, as seen in equation 2.22, the reverse current will increase due to the increase in temperature; this

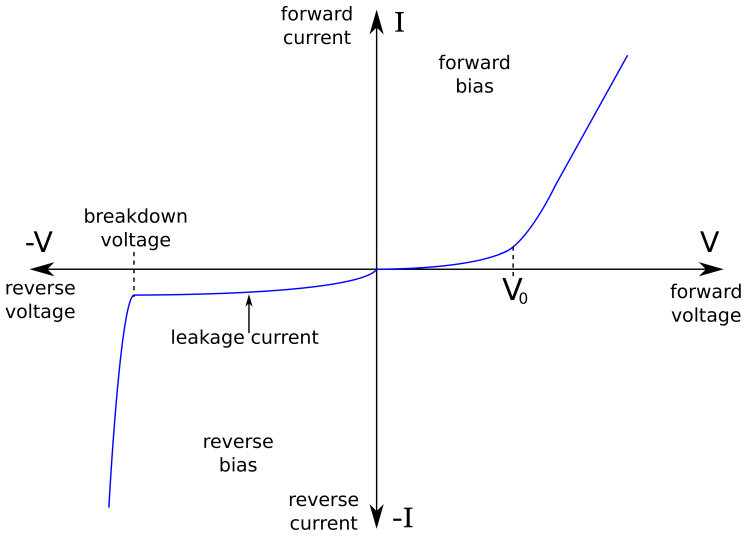


Figure 2.9: Current-voltage characteristics of a p-n junction in both forward and reverse bias.

creates a positive feedback loop and is called thermal runaway. This is pertinent to semiconductors with small band gaps such as Ge and it can be prevented by using a cooling system.

Zener breakdown occurs when a high enough field forces a current to flow through band-to-band tunnelling of carriers from the valence band to the conduction band. This happens because, for a reverse bias, the potential across the junction is large, causing the valence band of the p-type to reach the same level as the conduction band of the n-type. The movement of the valence electrons in the p-type is restricted by a triangular potential barrier whose height is given by the band gap. This can be overcome by the quantum mechanical tunnelling effect, creating a current. This sudden rise of reverse current destroys the depletion region. Zener breakdown occurs in highly doped p-n junctions were the resulting depletion region is small.

If the electric field is high enough, and the high field region is large enough, then breakdown due to avalanche multiplication will occur. In an electric field the free carriers are accelerated and gain kinetic energy. If their energy is high enough between collisions with lattice sites, they can knock out a bound electron (in the valence band) and promote it to the conduction band, creating an electron-hole pair. This process is called impact ionisation. The newly generated carriers can also acquire enough energy to liberate more carriers creating an avalanche effect, leading to a large increase in current and thus breakdown. This type of breakdown occurs in lowly doped p-n junctions that have a large depletion region.

2.1.9 Equations of state

In the previous sections the treatment of the primary action of the carriers through recombination-generation, drift and diffusion has been considered separately. By combining these, the overall description of a dynamic state in the semiconductor can be obtained, allowing the system variables to be determined under arbitrary conditions in device analysis.

The total current that flows inside a semiconductor takes into account the actions by both forms of carrier transport, drift and diffusion. By summing these together, the total current density for electrons and holes, respectively, is:

$$\mathbf{J}_{\mathrm{n}} = \mathbf{J}_{\mathrm{n,drift}} + \mathbf{J}_{\mathrm{n,diff}} = q\mu_n n\mathscr{E} + D_n \nabla n\,, \tag{2.23}$$

$$\mathbf{J}_{\mathrm{p}} = \mathbf{J}_{\mathrm{p,drift}} + \mathbf{J}_{\mathrm{p,diff}} = q\mu_p p\mathscr{E} - D_p \nabla p\,. \tag{2.24}$$

Hence the total current flowing through the semiconductor is:

$$\mathbf{J} = \mathbf{J}_{\mathrm{n}} + \mathbf{J}_{\mathrm{p}} \tag{2.25}$$

All the processes described in the previous sections can change the carrier concentration. The rate of change in the carrier concentration is given by the difference between the generation and recombination rates and the total flux through a surface, the total current density. All these effects can be combined to produce the continuity equations:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot \mathbf{J}_n - R_n + G_n\,, \tag{2.26}$$

$$\frac{\partial p}{\partial t} = \frac{1}{q}\nabla \cdot \mathbf{J}_p - R_p + G_p\,, \tag{2.27}$$

where R is the recombination rate and G is the generation rate. The electric field is calculated using the Poisson equation and the charge distribution:

$$\nabla \cdot \mathscr{E} = \frac{\rho}{\varepsilon_r \varepsilon_0}\,, \qquad \text{with } \rho = (p - n + N_D - N_A)^{1}\,. \tag{2.28}$$

By using all the aforementioned equations, semiconductor devices and their properties can be simulated. An exact solution is not possible due to the large number of degrees of freedom. Therefore, numerical methods are employed using iterative techniques carried out by a computer; this is explained in more detail in chapter 5.

2.2 Silicon as a particle detector

Due to the very well understood properties and the ability to manipulate these to produce a desired outcome, silicon is a very suitable choice for a detector of ionising radiation. Such properties include: a small band gap, leading to a large number of carriers per unit energy loss; mechanical stability, allowing for complex structures; a high material density, so that a large amount of charge can be produced per length of the particle track; and a high mobility of carriers, as a result the signal can be collected quickly. Another advantage of silicon is its wide use within the electronics industry. This means it is a mature technology that has advanced processing technology available which can be exploited for use in particle physics.

[1] note: in this equation ρ is the charge density and not the resistivity.

The main principle of a silicon detector is a segmented diode, created by using a p-n junction. One example of this is a p-in-n sensor where the substrate is n-type and is implanted at the topside with a p-type to produce a collection diode for holes. The sensor is then reversely biased from the backside through an additional n+ doped contact at the back, creating a depletion region that grows from the p-implant. The extra n+ implant ensures it is an Ohmic contact. When a particle traverses the sensor, it loses energy through ionisation and generates electron-hole pairs, as shown in figure 2.10. These are then mainly transported to the collection diode by drift, possibly with some contributions from diffusion; this depends on the operation conditions and the type of sensor. Once the charge is collected, the signal is then read out and sent to a chip that can process and interpret the signal.

One type of segmented detector is the strip detector, with typical pitches of 50–100 µm. These give both an energy and a position measurement. These involve using diodes to create thin parallel strips in the sensor with the position of the ionising particle given by the location of the strip that collected the signal. The readout of these strips is done at the periphery of the device. In order to get two-dimensional information, the backside can be implanted with strips rotated by an angle. This makes use of the fact that the two forms of carriers are swept in opposite directions by the electric field and provides twice the information per track. Another method would be to put two sensors back to back, again having a rotation angle between the two sets of strips. Double-sided sensors are

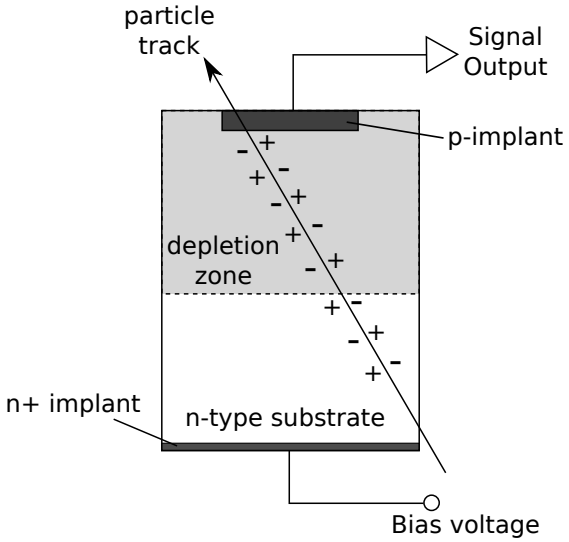


Figure 2.10: An illustration of a typical silicon detector based on the diode structure. The n-type substrate is implanted with a p-type to create a collecting electrode that is reversely biased from the back through an n+ implant. When a particle traverses the detector, electron-hole pairs are created and read out.

more difficult to produce, have complicated readout and in both cases, the information becomes obscured by ambiguous combinations of multiple hits for large occupancies. A way to overcome these problems is to segment the sensor into pixels which are the focus of this thesis.

2.2.1 Hybrid pixel detectors

The pixel detector is an extension of the idea of a strip sensor by incorporating an extra level of segmentation of the strips into pixels, usually forming square or rectangle shapes. As a consequence, the readout also follows the same segmentation and is done on a one-to-one basis, meaning that for each pixel on the sensor chip, there must be a pixel on the readout chip. If the pixels are very large and the number of channels is small this can be done by routing the signal to the periphery via wires. However, in particle physics applications these pixels tend to be small, ranging from a pitch of tens to hundreds of microns, and require the readout chips to be mounted directly on the sensor via flip-chip techniques, illustrated in figure 2.11. This adds extra restrictions to the detector: the readout channels must match the size of the sensor pixels and the readout chip must be very close (10–20 µm). Another requirement is that each pixel on the sensor side must be coupled to an amplifier on the readout side. This has so far been achieved using bump-bonding technology. This is where small solder balls provide an ohmic contact between the output pad of the sensor to the input pad of the readout chip.

Function in particle physics

The use of hybrid pixel detectors in high energy physics emerged when vertex detectors were needed to study short-lived particles through lifetime measurements such as the b-quark and tau lepton. This was previously done with strip detectors but they could

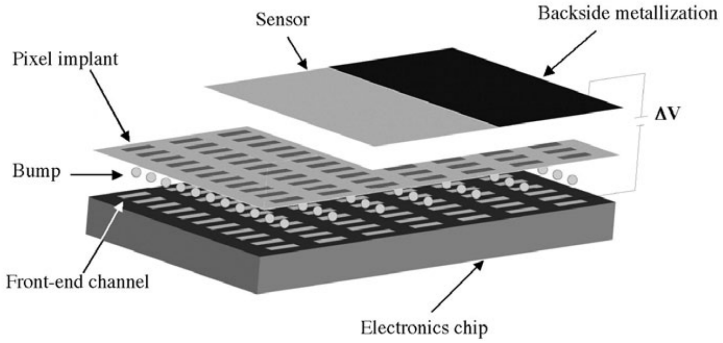


Figure 2.11: A schematic of a hybrid pixel detector showing the sensor bump-bonded to the readout chip on a one-to-one basis [12].

no longer cope with the high density of particles produced per collision at high energy colliders. Pixel detectors have increased spatial resolution as well as providing enough channels to operate in such an environment. The lifetime of these short-lived particles is in the picosecond range, after which the particles have moved a few mm away from the interaction point, the primary vertex PV. An illustration of such a decay topology is shown in figure 2.12. To measure the position of where the short-lived particle decayed, the secondary vertex SV, a spatial resolution with a precision of a few tens of microns is required [14]. This increases the ability to separate tracks from the primary and secondary vertices, allowing for the identification of charm and bottom hadrons. To reduce the ambiguity associated with this measurement, the detectors need to be placed as close as possible to the interaction point to minimise extrapolation errors. Layers of detectors are implemented to get multiple measurement points along the particle track. This increases the vertex resolution but introduces the negative effect of multiple scattering, equation 2.14, which can spoil the measurements. To reduce this the material in the detector is kept to a minimum, making thin sensors desirable. A magnetic field can also be applied to increase the vertex resolution.

Leakage current

The properties of the sensor can be derived from the semiconductor theory outlined earlier in this chapter. They give an indication on the quality of the sensor as well as the operating conditions needed. Since all pixel detectors are operated in the reverse bias mode, the leakage current comes into play. This is undesirable and action is taken in order to minimise it, as it is a source of noise and power consumption. For voltages below the full depletion, the leakage current is described by equation 2.21 and is roughly proportional to the square root of the bias voltage and the square root of the

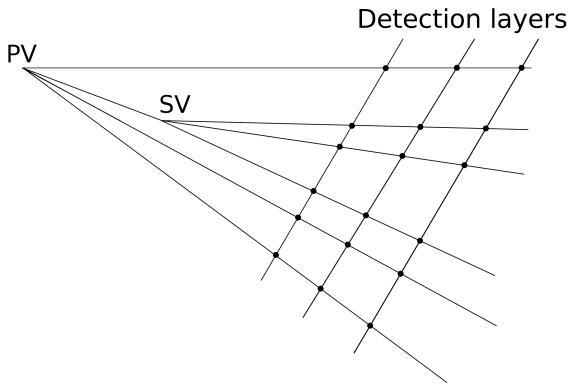


Figure 2.12: A schematic of the decay topology for a short-lived particle showing the primary vertex, PV, and the vertex of the short-lived particle, the secondary vertex, SV.

resistivity. Once full depletion is reached, the leakage current plateaus. If the applied bias is high enough then the sensor current will break down, potentially causing irreversible damage to the sensor especially for breakdown due to thermal runaway. Often the bias at which breakdown occurs, the breakdown voltage, determines the maximum operating voltage of the sensor; it can be determined by measuring the current as a function of applied voltage.

Depletion region

Often in particle detectors the diode junction occurs between one highly doped implant ($N > 10^{18}\text{cm}^{-3}$) and a low doped bulk ($N \approx 10^{12}\text{cm}^{-3}$). In addition to this the built-in voltage is typically much lower than the operation voltage and hence can be neglected. As a result equation 2.9 can be approximated to:

$$w \approx \sqrt{\frac{2\varepsilon_r\varepsilon_0 V}{qN}} = \sqrt{2\varepsilon_r\varepsilon_0\rho\mu V}\,. \tag{2.29}$$

N represents either N_A (p-type) or N_D (n-type) depending on the doping type. If a large enough reverse bias is applied, the depletion region can extend across the whole device; this voltage is called the full depletion voltage. For a device of thickness d the full depletion voltage, V_{dep}, is given by:

$$V_{dep} = \frac{qNd^2}{2\varepsilon_r\varepsilon_0}\,. \tag{2.30}$$

At this point, the sensor will collect charge created anywhere within this volume and this is therefore the standard voltage to operate at. However, the depletion voltage may be larger than the breakdown voltage so some sensors are operated below their depletion voltage, sometimes by design or due to damage caused by radiation.

Capacitance

Knowledge of the capacitance is important for the noise characterisation of the device. The more capacitance a pixel has the more noise the pre-amplifier will have in the readout chip; hence the goal is to minimise all capacitances. The capacitance of the sensor is made up of several parts: the capacitance of the pixels to the backside or bulk (C_b), the capacitance to the neighbour pixels (C_i) and capacitance to the readout chip and bump-bonds (C_c), shown in figure 2.13. The capacitance to the backplane or bulk can be estimated using the parallel plate capacitor estimation and is given by equation 2.10. Capacitance to neighbours, interpixel capacitance, can lead to cross-talk. This

is where signal on one pixel induces a signal to the neighbour, leading to an increase in the number of multi-pixel hits, possibly spoiling the position reconstruction. Other sources of capacitance are often small enough that they are dominated by the bulk and interpixel capacitances. The interpixel capacitance depends on the geometry, depth and size of the pixels. For rectangular pixels it is large and is often the dominant source of capacitance. However, for square pixels it is smaller. In applications for the vertex detector, the pixels are very small and are square shaped meaning that the dominant source of capacitance is the contribution from the bulk.

Spatial resolution

A key purpose of pixel detectors is to measure the position of a track. The main factor in determining the spatial resolution is the pitch p of the pixels and the method of readout. For a readout system that uses binary information the measurement precision, defined as the root-mean-square distance of the measured coordinate from the true hit position [12], is given by:

$$\sigma_x = \frac{p}{\sqrt{12}}. \tag{2.31}$$

This applies when only binary information is used, if a pixel is hit the position of the track is taken as the centre of the pixel. This is used when no energy information is required, there is no charge sharing between pixels and the precision given by the pitch is sufficient.

To improve on this, and to also incorporate energy measurements, analogue information about the hit, such as the pulse height measurement, can be used. This produces a signal that is proportional to the collected charge. An improvement can only be made if the charge produced by the particle is shared between multiple pixels; if it only cre-

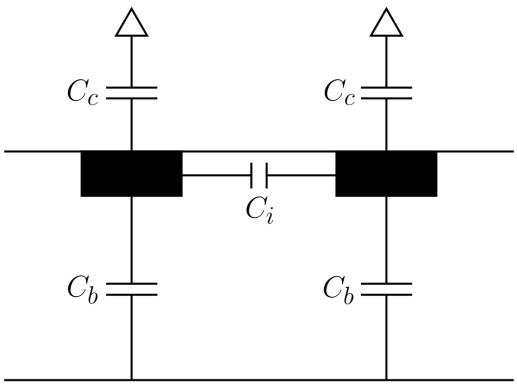


Figure 2.13: A sketch of the dominant capacitances in a pixel detector.

ates signal in one pixel, then the precision is still given by equation 2.31. Therefore, effort is sometimes put into increasing charge sharing in the sensor. This can be done by diffusion of charge, angled tracks, choosing a smaller pitch and by Lorentz angle (movement of carriers in the sensor due to a magnetic field). Charge interpolation methods are then used to find the coordinate of the particle, such as the centre of gravity of the signal, which will be discussed in chapter 7.2.1. In this case, the precision is limited by the signal-to-noise ratio S/N of the sensor [14]:

$$\sigma_x = \frac{p}{S/N}\,. \tag{2.32}$$

This assumes that the charge distribution is linear, which is often not the case, hence more complicated methods are used such as the η-algorithm, outlined in chapter 7.2.1. For charge sharing due to inclined tracks, the limitation of the measurement precision arises from the Landau fluctuations of deposited charge. Overall, there are many factors that influence the spatial resolution obtained by the sensor: the pixel pitch, choice of readout, the amount of charge sharing and the position finding algorithms.

2.2.2 Monolithic detectors

Another way to produce pixel detectors is to incorporate the readout electronics within the pixel, the monolithic active pixel sensors (MAPS). This can offer advantages in terms of cost, material budget and detector capacitance. They are primarily designed to meet the two major requirements of most modern high energy physics experiments: to be radiant tolerant and have a low power consumption. One such current application of this type of MAPS detectors in high energy physics is for the vertex detector at STAR [15]. The current design for MAPS is to have the circuitry embedded in a low resistivity substrate and for the majority of the charge collection to occur via diffusion, which is slow (figure 2.14(a)). Moreover, the in-pixel circuitry is limited resulting in simple circuits that do not have a fast clock in each pixel, leading to long readout integration times. One advantage is that less material is used as the readout chip and sensor are combined.

Recent developments have introduced full CMOS circuitry into the pixel as well as using depleted MAPS to combat these short comings. This is done by adding an additional deep p-well implant into the sensor to shield the electronics from the substrate (figure 2.14(b)). To achieve a larger depletion region, either high-voltage CMOS (HV-CMOS) or high-resistivity CMOS (HR-CMOS) is implemented. One application for these type of sensors is for the ALICE ITS upgrade [16]. However, even with the more complex

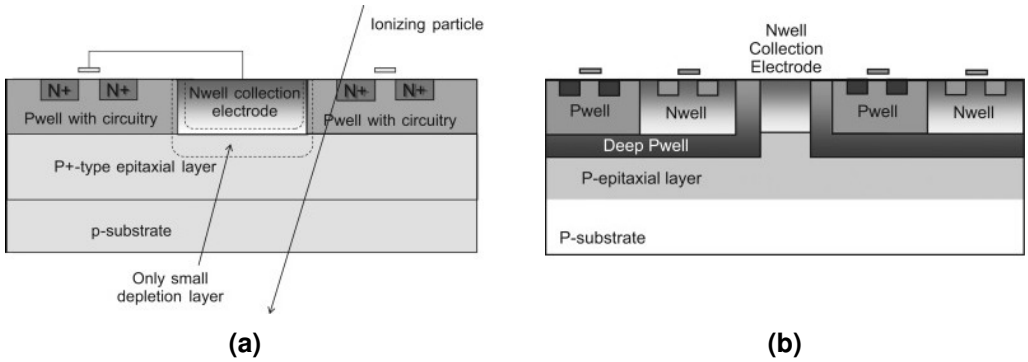


Figure 2.14: Cross-section of two types of MAPS detectors: (a) traditional low resistivity and (b) high resistivity with a deep p-well implant for shielding [18].

circuitry available, the electronics are still limited in comparison to dedicated readout chips used in hybrid detectors. This is especially the case for very small pixel sizes ($< 50\,\mu m$) due to the rather large feature size of the monolithic processes currently available ($< 100\,nm$). This is what motivated the design of active HV-CMOS sensors capacitively coupled to high performance readout chips [17].

2.3 Readout electronics

The front-end readout electronics that are located on a readout chip can be the largest challenge in the design of a pixel detector system. Their function is to process the signal that the sensor has collected and transport it out of the detector to be recorded. The difficulty comes from providing amplification, hit discrimination and readout for thousands of pixel circuits while simultaneously having low noise, low power consumption and being fast. For most sensor technologies the signal generated (typically 5000-30000 e^-) is quite small so some form of amplification is needed before it can be processed. This is typically done on the readout chip, or in amplifiers integrated into the sensor.

The interplay between the sensor and the electronics is very important when designing a detector system as it impacts on the performance such as the resolution and noise. The operational performance of the electronics is mainly determined by the signal-to-noise ratio. Acceptable limits for this are dictated by three factors: efficiency, noise occupancy and the resolution required. The threshold is a parameter used to decide whether or not a hit has occurred. Having a large threshold means that some tracks may not produce hits, reducing the efficiency and the number of multi-pixel clusters, thus causing an increase in spatial resolution. On the other hand, having a low threshold means that fluctuations due to noise could produce unwanted hits, producing a large noise occupancy, which can spoil the spatial resolution and possibly

saturate the data acquisition system. Typical values for a pixel readout chip are a noise of $< 100\,\mathrm{e}^-$ and an operation threshold which is $\gtrsim 5\sigma$ of the noise distribution above the baseline. The baseline is the threshold value at which half of the signals generated are recorded, i.e. for 100 signals the pixel records 50 hits. A global baseline is the mean of the baselines for all pixels.

In modern particle detectors the speed and the power consumption are also important factors that are strongly determined by the readout electronics. The high signal rates require a fast readout in order to avoid pile-up, but faster electronics result in higher noise and power consumption. The vast number of channels in a very small volume means that power consumption becomes a major problem and action is taken to reduce this as much as possible. Therefore, the design needs to take into account and optimise the speed, noise and power consumption. A simplified view of the components that go into the readout chip for hit detection, arrival time and energy measurements, that is typically used in readout chips of pixel detectors, is shown in figure 2.15.

Charge sensitive amplifier

A charge sensitive amplifier (CSA) is an integrator that takes current as input and converts it to a voltage with some gain factor. This effectively acts as a measurement of the input charge. These are used because of their low noise, stability and the fact that the signal of silicon sensors is charge. A CSA contains: an amplifier, a feedback capacitor, C_f, (which is between the input and output of the CSA) and a reset circuit that restores the voltage back to the baseline; the circuit is shown in figure 2.15. When a charge is

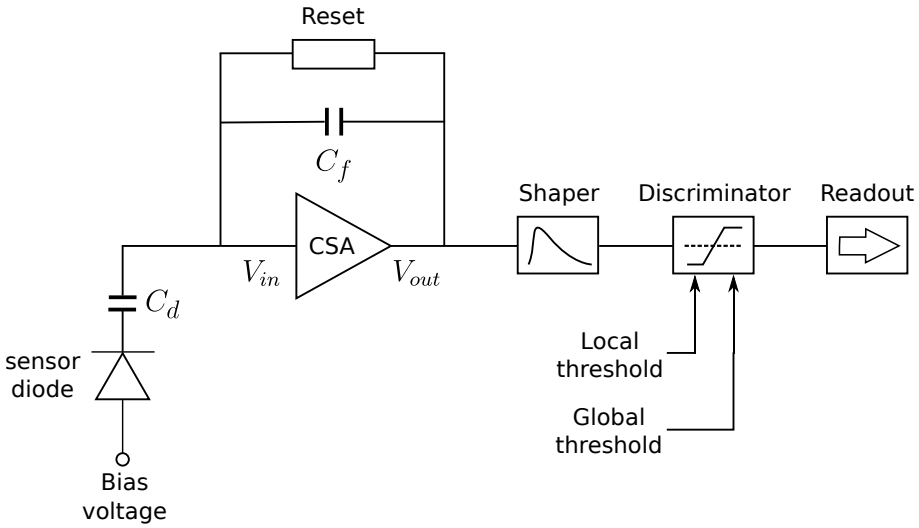


Figure 2.15: The basic components of a pixel readout chip showing the sensor diode, where C_d is the detector capacitance and C_f is the feedback capacitance of the charge sensitive amplifier (CSA).

induced on the sensor it charges the sensor capacitance, C_d, producing a step in the output voltage of the amplifier. The gain a of this type of amplifier is:

$$\frac{\mathrm{d}V_{out}}{\mathrm{d}V_{in}} = -a\,, \tag{2.33}$$

where $V_{out/in}$ is the output/input voltage of the CSA [14]. Leading to an output voltage per unit input charge Q_{in} of [14]:

$$A = \frac{\mathrm{d}V_{out}}{\mathrm{d}Q_{in}} = \frac{1}{C_f + \frac{C_f + C_d}{a}} \approx \frac{1}{C_f}\,, \quad (\text{for } a \gg 1)\,. \tag{2.34}$$

Shaper

The first component after a CSA is the shaper or filter. This takes the output pulses and turns them into signal shapes that are easier to handle. The reason this is needed is that the output pulses of the CSA have very long tails that can be detrimental in high rate environments as the pulses may overlap and may thus be hard to distinguish. Complementary to this, the shaper also filters much of the noise improving the signal-to-noise ratio considerably [14].

Discriminator

The next stage in the signal processing is the discriminator which compares the now shaped pulse to a threshold value. If it is higher, then a hit is deemed to have occurred, producing a binary signal. The threshold value is universal over the whole pixel matrix. Due to several effects such as the variation in the fabrication, the threshold on a pixel-to-pixel level has some discrepancy causing a reduced efficiency or increased noise in the pixel relative to the other pixels. To compensate for this a local threshold tuning is applied to get a uniform response across the pixel matrix. The threshold in the discriminator is a voltage and is set using a DAC (digital-to-analogue converter) value.

The readout chip has a clock that cycles at a certain frequency and increments a counter each cycle within a given shutter period. The time at which the signal first crosses the threshold is called the time-of-arrival (ToA). To obtain analogue information of the charge, the width of the discriminator output pulse is measured because it is proportional to the input charge and depends on the reset circuit. This is called a time-over-threshold (ToT) measurement, shown in figure 2.16. In this case the readout chip clock is used to measure the amount of times that the hit stays above the threshold.

The ToT measurement converts the analogue signal in the discriminator to a digital signal, so it has units of ADC (analogue-to-digital converter). The measurement of charge can be used to improve the spatial resolution of the sensor and correct for time walk effects. Time walk is the variation of the time at which the measured signal pulse, in the discriminator, crosses the threshold and depends on the signal amplitude. This results in a broad time distribution.

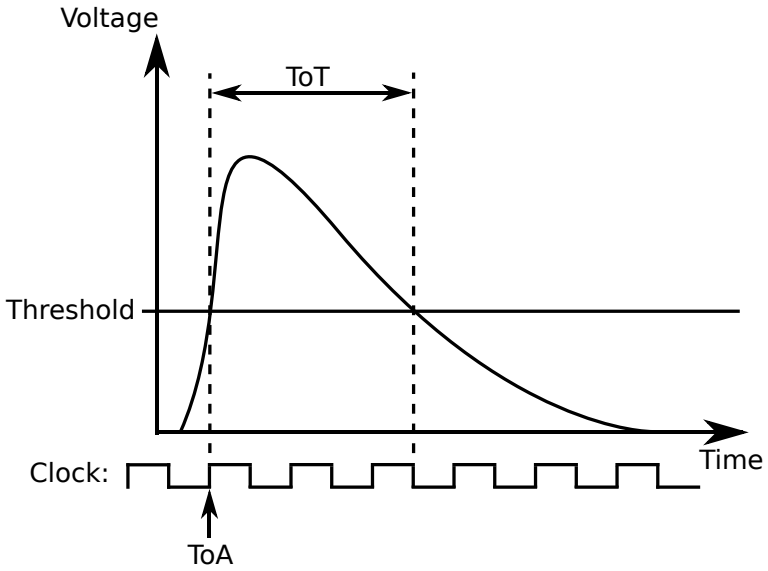
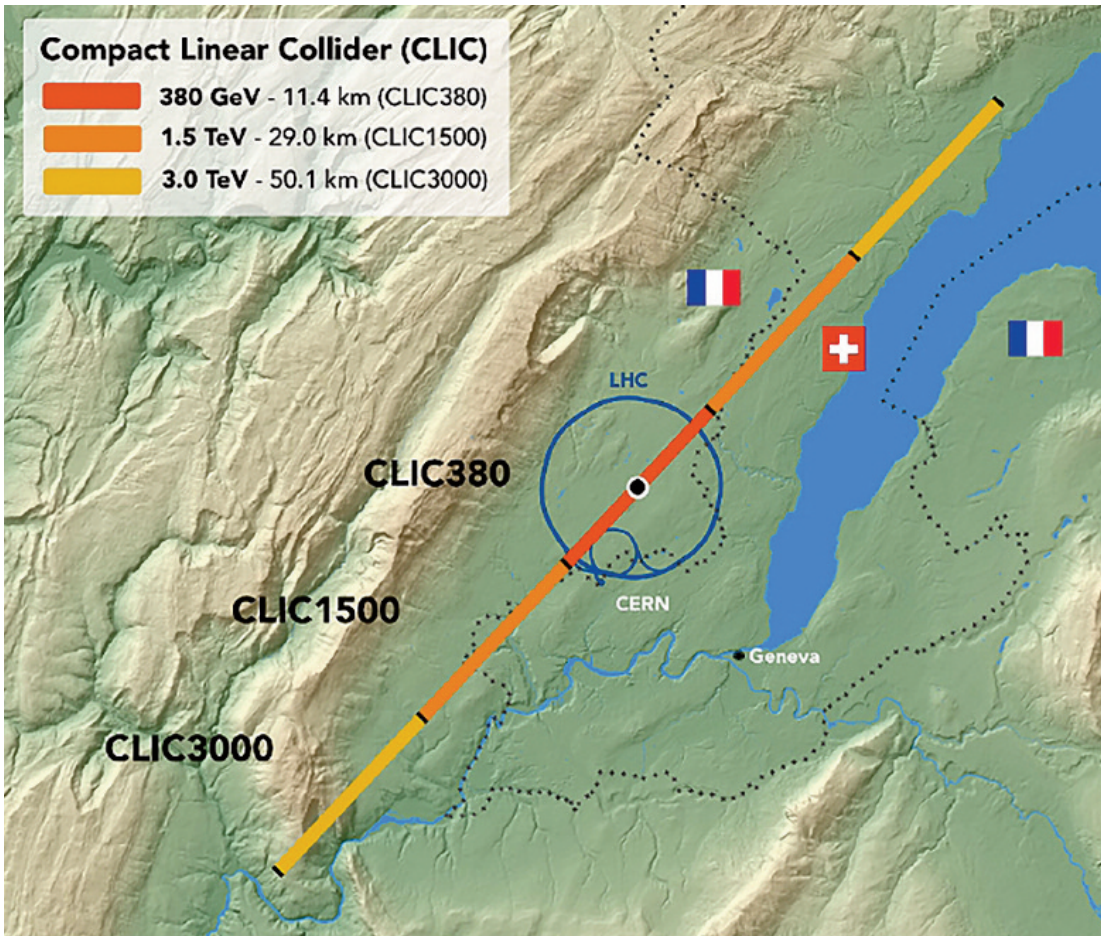


Figure 2.16: An example of a pulse input to the discriminator and the resulting time of arrival (ToA) and time over threshold (ToT) measurements. When the pulse is above the threshold value the ToT measurement begins. The frequency of the clock counter increase determines the precision of the ToA and ToT measurements.



An overview of the footprint of CLIC for the proposed three energy stages showing the surrounding area of CERN and Geneva. (see page 37)

CHAPTER 3
CLIC overview

The aim of this chapter is to give the reader a summary of the main components of the compact linear collider (CLIC) and is divided into four sections. The first section will focus on the physics motivation for CLIC and the impact this has on the accelerator and detector. After this, the accelerator will be described giving details of the technology and the foreseen stages of CLIC. Then a general overview of the detector concept will be given providing a breakdown of the constituent parts. Finally, a summary of the physics measurements that are expected to be carried out will be given.

3.1 Physics motivation

The goal of CLIC is to provide a wide range of precision physics measurements and sensitivity to new physics phenomena available at the multi TeV energy frontier, taking advantage of the relatively clean environment found at a lepton collider compared to a hadron collider.

The importance of the physics at CLIC is twofold, not only is the potential to find new physics important but also the precision measurements, which will play a key role in developing our knowledge of particle physics. CLIC is envisaged to address the fundamental questions in the post-LHC era through precise mass and coupling measurements to the Higgs boson and top quark. The programme for CLIC will extend over a broad range of physics including, in particular, measurements of the Higgs boson and top-quark. CLIC will also conduct direct and indirect searches for beyond standard model (BSM) physics.

Based on the physics goals of CLIC, requirements for both the accelerator and detector are generated. The choice of colliding electrons and their anti-matter partner, positrons, in the multi-TeV energy range means that the collider will need to be linear. The reason for choosing a linear collider is that when charged particles move in a curved path they emit synchrotron radiation in the form of photons, thus loosing energy and limiting the final collision energy. For a circular collider this is unavoidable, so one way to overcome this is to accelerate the electrons in a straight line. As a result, large acceleration gradients are needed to reach the desired centre-of-mass energies, for CLIC this is $100\,\mathrm{MV\,m^{-1}}$. To achieve this, novel acceleration schemes need to be designed and tested; for CLIC the two-beam acceleration scheme is chosen, providing colliding beams up to 3 TeV. Another consideration for the accelerator is to polarise the electron beam at an expected value of 80%, as this greatly enhances the physics potential [19].

The reason for this is that it increases the probability of certain physics interactions, suppresses background rates and provides additional information about the produced particles such as quantum numbers and helicity.

The breadth of the CLIC programme will be extended by operating CLIC at different centre-of-mass energies, the values of which have been optimised, to create the largest possible physics reach [20–23]. The energy stages have been based on current knowledge and are subject to change depending on any shifts in the physics landscape over the next years such as, for example, a new discovery by the LHC.

3.2 CLIC accelerator

An extensive design study on the CLIC accelerator concept has been carried out and has successfully addressed the main technical challenges, details of which are found in the conceptual design report (CDR) [5]. Unlike traditional accelerators, which are based on radio frequency (RF) cavities called klystrons, CLIC proposes to use a novel two-beam acceleration scheme to collide electrons and positrons. The centre-of-mass energies these particles will collide at will be ramped up in three stages: 380 GeV, 1.5 TeV and 3 TeV, with a peak luminosity of $2 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at 3 TeV [23]. The peak luminosity is defined as the luminosity within the 1% centre-of-mass energy peak. The key parameters that describe the various stages are shown in table 3.1.

Table 3.1: The key parameters of the various stages for CLIC, from [23].

Parameter	Symbol	Unit	Stage 1	Stage 2	Stage 3
Centre-of-mass energy	$\sqrt{s}$	GeV	380	1500	3000
Repetition frequency	f_{rep}	Hz	50	50	50
Number of bunches per train	n_b		352	312	312
Bunch separation	Δt	ns	0.5	0.5	0.5
Accelerating gradient	G	$\mathrm{MV\,m^{-1}}$	72	72/100	72/100
Total luminosity	$\mathscr{L}$	$10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$	1.5	3.7	5.9
Luminosity above 99% of $\sqrt{s}$	$\mathscr{L}_{0.001}$	$10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$	0.9	1.4	2
Main tunnel length		km	11.4	29.0	50.1
Number of particles per bunch	N	10^9	5.2	3.7	3.7
Bunch length	σ_z	µm	70	44	44
IP beam size	σ_x/σ_y	nm	149/2.9	∼60/1.5	∼40/1
Estimated power consumption	P_{wall}	MW	252	364	589

3.2.1 Two-beam acceleration

The motivation behind the two-beam acceleration scheme is that, in order to achieve multi-TeV energies in a practical length, the beams need to have a high acceleration gradient of 100 MV m^{-1}. Contrary to circular colliders, where the beams are sent around multiple times gaining an energy kick for each rotation, in a linear collider the particles only pass through once. The 100 MV m^{-1} acceleration gradient cannot be achieved with the current superconducting klystron technology. Thus, an alternative approach is needed.

The principle of the two-beam acceleration scheme is to transfer energy for the main beam from a secondary beam, the drive beam. This is an intense beam that is accelerated using klystrons and runs parallel to the main beam, with a separation distance of 60 cm. It is decelerated and has its power extracted and turned into RF power by the power extraction transfer structures (PETS) that feed two accelerating structures, shown in figure 3.1. The RF power is transferred to the main beam via waveguides. The drive beam has a very high current (100 A), a high density of electrons and a relatively high energy (2.4 GeV). So when about 84% [5] of their energy is transferred to a low current (1.2 A), less dense beam, it produces a large acceleration.

3.2.2 CLIC beam structure

The bunches of electrons or positrons that are collected from the source contain 10^9 particles and are separated by 0.5 ns. The bunches are gathered together in what are called bunch trains. Each train consists of 312 bunches and is separated from the next bunch train by a gap of 20 ms, producing a repetition frequency of 50 Hz.

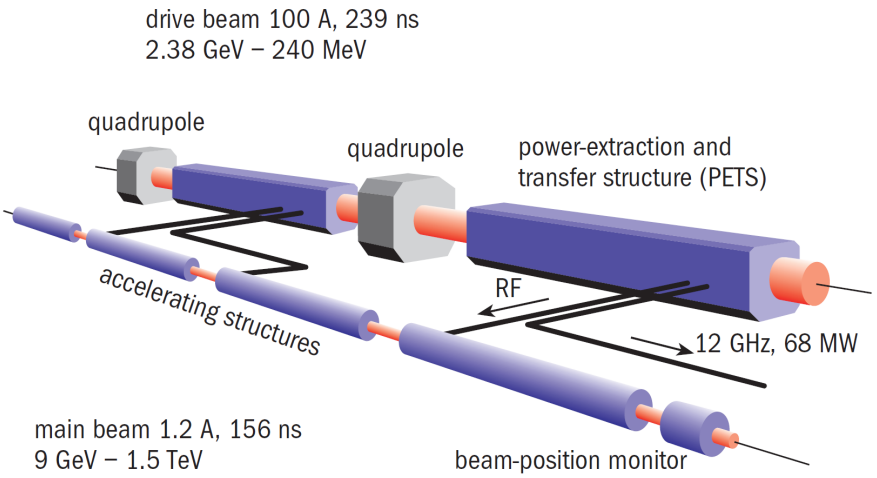


Figure 3.1: An illustration of the two-beam acceleration scheme. The drive beam is decelerated and has its energy transferred by the PETS to the main beam [5].

The exact amount of spacing between the trains is determined by the long-range wakefields that induce disruptive electromagnetic fields. These must be removed from the accelerating structure before the next bunch arrives, which currently can be done at 0.5 ns. In order to minimise losses, the trains need to be as close to each other as possible. To avoid the beams colliding at any place other than the interaction point (IP), a crossing angle is introduced. However, the angle has a detrimental effect on the luminosity achieved. To balance this an angle of 20 mrad is chosen to minimise the loss to below 1% [5].

3.2.3 CLIC accelerator complex

The layout of CLIC showing the accelerator concept and all the constituent parts that make up the accelerator complex is shown in figure 3.2. One of the two beams is called the drive beam. It starts off with acceleration by traditional klystrons and then enters the delay loop and two combiner rings (CR). These are used to build up the beam intensity to generate a large current of 100 A. The drive beam is then transported to the main linac tunnel and each bunch train will undergo deceleration in one of the 25 two-beam sectors. Once it has undergone deceleration, the drive beam is turned away from the main linac and then sent to a beam dump, one for each sector.

The generation of the two main beams occurs in a different place to the drive beam and starts off in the injectors, with the electrons having a polarisation. Following this they are then boosted by a common linac and transported to their respective main beam lines ready for the two-beam acceleration. Once there they undergo an acceleration of 100 MV m^{-1}, passing through the 25 sectors, each of length 878 m and containing

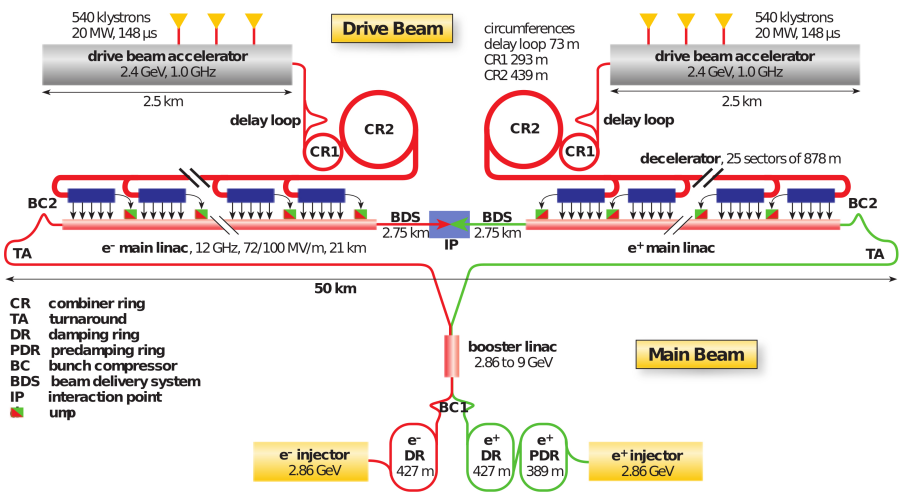
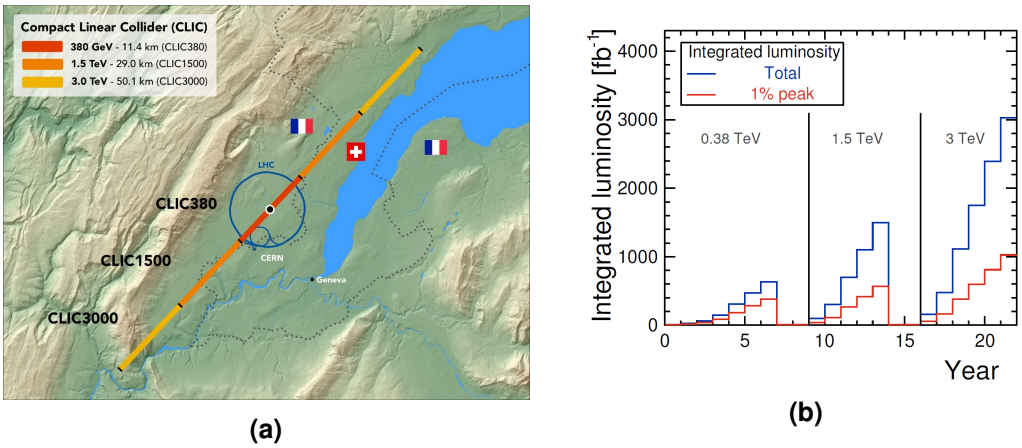


Figure 3.2: The CLIC accelerator concept at 3 TeV, from [23].

3000 accelerating structures; each providing the main beam with 62 GeV of energy. After passing through all the two-beam structures, they go to the beam delivery section (BDS) and then collide at the interaction point.

3.2.4 CLIC staging baseline

Various studies of the physics at CLIC have indicated that implementing a staged baseline scenario will maximise the physics potential, permitting a comprehensive programme that will span several decades [23]. Performance, cost and power optimisation studies have been conducted on the accelerator to determine whether the staged scenario is achievable [5, 23, 24]. The current foreseen stages will deliver collisions at the centre-of-mass energies: 380 GeV, 1.5 TeV and 3 TeV, for which the corresponding layout is shown in figure 3.3(a). Overall the CLIC project is foreseen to last 22 years from the start of operation and will have a 7 year construction period [24]. The operational period is split up into 7, 5 and 6 years for the three energy stages (from 380 GeV to 3 TeV) with upgrade periods lasting 2 years. During this time the expected integrated luminosity is shown in figure 3.3(b). Each stage has a target for the integrated luminosity: 500 fb^{-1} at 380 GeV, 1.5 ab^{-1} at 1.5 GeV and 3 ab^{-1} at 3 TeV, which define the duration of the stage. On top of this, there will be a top threshold scan around 350 GeV producing an additional 100 fb^{-1} during the few months in which this will take place.



(a) (b)

Figure 3.3: (a) An overview of the footprint of CLIC for the proposed three energy stages showing the surrounding area of CERN and Geneva, from [25]. (b) The integrated luminosity for CLIC at the three staging scenarios. The time is taken from when the beam commissioning begins, from [23].

3.2.5 CLIC experimental environment

The parameters of the beam will determine the physics environment and impact heavily on the detector design. There are three main features of the CLIC machine. The first is the small beam size. This means there is a high bunch charge density and when the two beams come in range of each other's strong electromagnetic field, they radiate photons. This is called beamstrahlung. As a result, the colliding particles may not have the nominal centre-of-mass energy. Instead the collisions have a peak at $\sqrt{s}$ corresponding to collisions with no beamstrahlung and a long tail down to low energy values. This is why an important parameter for lepton colliders is the peak luminosity. The other main feature is the large number of beam induced backgrounds. One of these is the production of $e^+ e^-$ pairs, both coherent and incoherent, that arise from the beamstrahlung photons [26]. The separation is defined by the interaction of the photons with the collective electromagnetic field of the opposite beam, coherent, or with individual particles of the other beam, incoherent. Hadronic final states from two photon interactions can also be produced from the interaction of real and virtual photons from the beams [27].

The third feature of the CLIC machine is the beam structure, specifically the train repetition frequency of the beam, 50 Hz, and the bunch spacing, 0.5 ns. As a result, multiple bunch crossing will be selected in a DAQ cycle. This, along with the beam induced backgrounds means that all detectors will require good time resolution of a few ns, in order to be able to reject out-of-time background hits. As a result, for each bunch train, consisting of 312 bunches, there is a 156 ns active period. These are separated with a gap of 20 ms before the next train, during which the data is read out to the off-detector electronics. In this scenario there is no trigger selection in the data and as a consequence, the data will have to contain enough timing resolution in order for the offline analysis to separate it into signal from background. This is because the timing of the collisions is well defined and any events that occur outside this time window will be due to background. So, in order to maximise the potential to distinguish between the two without a trigger, a good timing resolution is needed.

3.3 CLIC detector

The size of the CLIC detector is 12.9 m in height and 11.4 m in length (figure 3.4). The principal concept that most modern physics detectors are based on is that of a series of layered sub-detectors that have specific jobs and work together to give an overall picture of what occurred. The CLIC detector has a cylindrical shape consisting of concentric layers with the interaction point in the centre. The layers, from the inside out, are as follows: the vertex detector (which will be described in detail in chapter

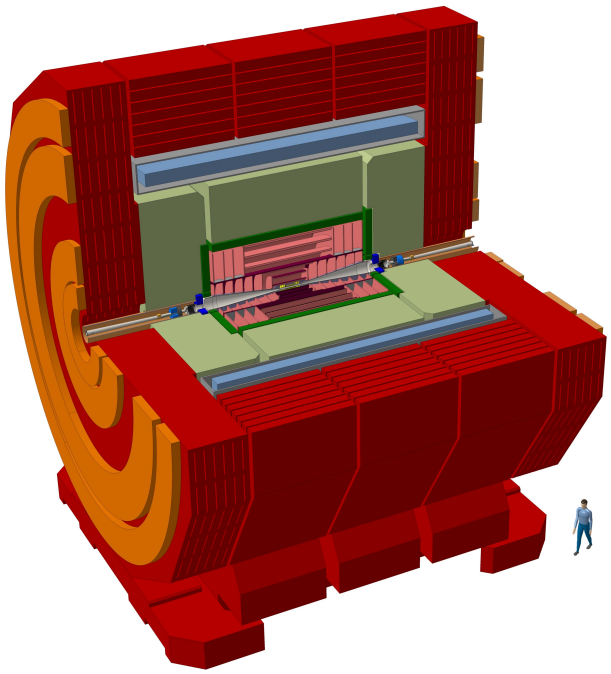


Figure 3.4: An illustration of the CLICdet model with dimensions of 12.9 m in height and 11.4 m in length [28].

4), the tracking detector, the electromagnetic calorimeter, the hadronic calorimeter, the solenoid magnet and the magnetic yoke with muon detection. There is also a need to have very forward detectors: the luminosity calorimeter (LumiCal) and the beam calorimeter (BeamCal) for beam monitoring. The main layers will be described in the following sections.

The design of all these sub-detectors will need to cope with the environmental conditions the accelerator will produce, on top of the physics requirements. The physics requirements are:

- impact parameter resolution (equation 4.1) for tracks of $a \approx 5\,\mu\text{m}$ and a multiple-scattering term of $b \approx 15\,\mu\text{mGeV}$
- track momentum resolution of $\sigma_{p_T}/p_T^2 = 2 \times 10^{-5}\,\text{GeV}^{-1}$
- jet energy resolution of $\sigma_E/E \lesssim 3.5$–5% for jet energies in the range 50 GeV to 1 TeV
- detector coverage down to very low angles,

more information can be found in [6].

3.3.1 Tracking system

The tracking system is in place to make precise momentum measurements of charged particles while being as thin and light as possible (1–2% X_0 per detection layer) in order to limit the amount of multiple scattering spoiling the measurement precision. It must be able to do this for a wide range of momenta at a very high efficiency over a large solid angle. To measure the track momentum information of the charged particle trajectory, a magnetic field is needed. The goal is to achieve a transverse momentum resolution of $\sigma_{p_T}/p_T^2 = 2 \times 10^{-5}\,\text{GeV}^{-1}$ for high momentum tracks [6]. Obtaining this value depends on optimising three factors that influence the track momentum resolution: the tracker radius, the magnetic field and the single point resolution of the tracker sensors. For CLICdet these are a radius of 1.5 m, a magnetic field of 4 T and a single point resolution in the $r\varphi$-plane of 7 µm.

The tracking requirements are very demanding for the high energies and high beam-induced background levels found at CLIC. Muon background from the beam delivery system along with beam induced backgrounds from incoherent pairs and $\gamma\gamma \rightarrow$ hadrons result in high occupancy levels. Moreover, the tracker readout system must have precise time binning (10 ns), more details about the readout can be found in [29].

The tracker will be all silicon with a radius of 1.5 m and a length of 4.4 m, the layout is shown in figure 3.5. A support tube, with an inner and outer radius of 0.575 m and 0.6 m, divides the tracker into two regions: the inner tracker and outer tracker. The tube was chosen to be a certain distance from the conical vacuum pipe to maximise the tracker coverage in the forward disks. The inner tracker comprises of 3 barrel layers with 7 disk layers either side. For the outer tracker there are also three barrel layers but it only has 4 disk layers. The total active area will be in the order of 100 m^2.

3.3.2 Calorimeters

A large impact on the overall design of the detector comes from the requirement on the jet energy resolution, of $\sigma_E/E \approx 3.5\%$. The reason for this is that many interesting physics processes will be characterised by multi-jet final states and it is important to precisely reconstruct the jets for event identification and event reconstruction. for example, the calorimeters will need to be able to distinguish between jets coming from W, Z and H. event reconstruction is especially difficult at high energies where the boosted topology of the Higgs production from W W fusion could lead to jet merging. High granularity particle flow calorimeters are foreseen to achieve this, as well as being able to cope with the backgrounds [30]. The principle is to track and identify every particle throughout the detector and reconstruct the jet energy as the sum of the individual particles. This would not be possible in a calorimeter with coarse sampling because

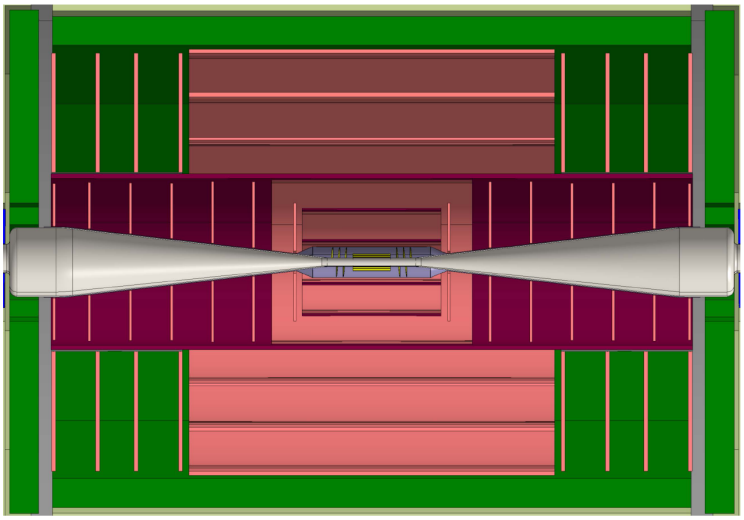


Figure 3.5: The layout of the CLICdet tracker [28].

the particles in a jet are very close together making it hard to reconstruct them individually. High granularity calorimeters are designed to do this. Another factor to consider for the calorimeters is the time stamping to facilitate the separation of signal from backgrounds. This results in the starting time of the shower needing to be determined with a time resolution of ≈ 1 ns [6]. Another feature is that the calorimeters need to be placed as close as possible to the tracker to be able to match the tracks, hence the calorimeters will be inside the solenoid magnet.

Electromagnetic calorimeter

The primary function of the ECAL is to identify and measure the energies of photons, electrons and positrons. On top of this, it has to deal with possible overlaps coming from charged particles and early showers from hadrons. To optimise the separation, a fine lateral segmentation is needed along with the fine granularity. Extensive studies on the optimal cell size and the number of layers have been carried out comparing different scenarios [28]. The technology foreseen for the ECAL is a tungsten absorber and silicon pad readout with 5×5 mm^2 cells and is under development by the CALICE collaboration [31]. The number of layers is set to 40 due to the improved energy resolution compared to using fewer layers, while still retaining a manageable mechanical complexity. The total thickness is around $22\,X_0$ [28] to ensure all the energy of the electromagnetic events are contained within the ECAL.

Hadronic calorimeter

A crucial requirement of the hadronic calorimeter is the time-stamping needed, because of the high background from two photon processes, and because the calorimeter

integrates over several bunch crossings. Another key factor to optimise is the jet energy resolution. Consequently, 60 layers of steel absorber plates with a cell size of $30 \times 30\,\mathrm{mm}^2$ are chosen. The detector cells are made of polystyrene scintillator tiles which are individually read out by SiPMs. This design is based on developments from the CALICE collaboration [32]. The total thickness of the full hadronic calorimeter is 7.5 interaction lengths [28].

Forward calorimeters

In the very forward regions two complementary calorimeters are included: the luminosity calorimeter (LumiCal) and the beam calorimeter (BeamCal), as illustrated in figure 3.6. The purpose of the LumiCal is to measure the luminosity via the number of Bhabha scattering events at low angles. The purpose of the BeamCal is to make fast estimates of the luminosity and the tagging of high energy electrons. Both detectors allow for coverage down to very low angles for the reconstruction of electrons and photons. The development of these calorimeters is done by the FCAL collaboration [33].

3.3.3 Detector magnet and muon system

The strength of the magnetic field will be 4 T and has been chosen so that the tracking requirements can be met. Outside the solenoid an iron return yoke is used to contain the field inside the detector as much as possible by reducing stray fields outside the detector. This also aids the muon momentum measurement by having a significant magnetic field inside the yoke. Inside the yoke it is 1.5 T and points in the opposite direction to the field inside the solenoid. Within this yoke is where the muon detectors reside; these are used to identify muons with a high efficiency, $> 90\%$ [6]. There will be 6 muon layers with an additional 7th layer in the barrel placed close to the solenoid. The technology proposed for these are resistive plate chambers that will have a cell size of $30 \times 30\,\mathrm{mm}^2$, with a possibility to use crossed scintillator bars [28].

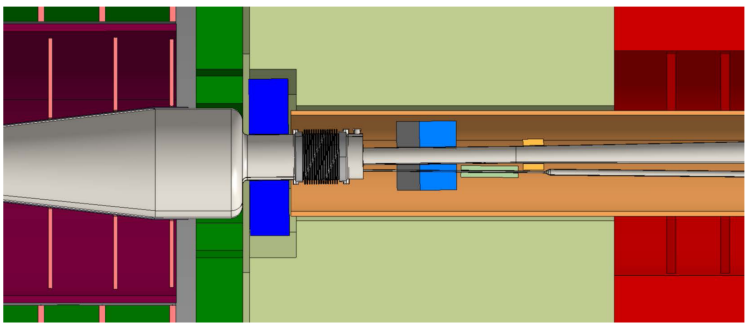


Figure 3.6: The CLICdet layout in the forward region as seen from the top. The LumiCal is shown in dark blue and the Beam Cal is in light blue [28].

3.4 Physics at CLIC

With CLIC being constructed in three energy stages there will be a focus on certain processes at each stage and the main measurements are as follows. For the first stage at 380 GeV, model independent measurements of the Higgs couplings to both fermions and bosons will be made through Higgsstrahlung and WW-fusion production processes. On top of this there will be precision top-quark measurements at this stage. The intermediate stage at 1.5 TeV will allow the measurement of additional Higgs and top-quark properties such as the top-Yukawa coupling, rare Higgs branching ratios and Higgs self-coupling. It will also provide the possibility to look into new physics phenomena beyond the SM. The final stage of 3 TeV is expected to supply the best sensitivity to new physics via indirect searches. It will also allow the possibility to study direct production. Pair-produced particles with masses up to 1.5 TeV and single particles with masses up to 3 TeV will be potentially discovered and accurately measured, e.g. dark matter candidates. A breakdown of the physics programme into two key areas is detailed below.

3.4.1 Higgs physics

With the discovery of the Higgs boson at the LHC [1, 2], the last piece of the puzzle for the standard model (SM) was put in place. However, many questions about the exact nature of this particle are still to be answered. It could be a fundamental boson of the SM or part of an extended Higgs sector, such as the five scalar particles found in the supersymmetric extension of the SM. A comprehensive study of this particle and its properties is needed and the clean $e^+ e^-$ environment found at CLIC can deliver this to a high precision. The key properties to be studied are the Higgs couplings.

The development of the leading-order $e^+ e^-$ SM Higgs production cross sections with the centre-of-mass energy is shown in figure 3.7 for a Higgs mass of 126 GeV. The three highest cross section Higgs production modes are: the Higgsstrahlung process ($e^+ e^- \rightarrow H Z$), the WW-fusion process ($e^+ e^- \rightarrow H \nu_e \overline{\nu}_e$) and the ZZ-fusion process ($e^+ e^- \rightarrow H e^+ e^-$). The lowest order Feynman diagrams for these are shown in figure 3.8. The lower energy stage is mainly motivated by the measurement of the Higgs coupling g_{HZZ} in a model-independent way, accessible via the Higgsstrahlung process. This model-independent measurement is unique to electron-positron colliders.

The large cross section of the WW-fusion process in combination with the high luminosities at CLIC mean that there will be large signal samples available for analysis. This allows for measurements of the Higgs couplings to both fermions and gauge bosons with great precision $\mathcal{O}(1\%)$ [20]. Also accessible, at the higher centre-of-mass energies of 1.5 TeV and 3 TeV, are rarer processes sensitive to the top-Yukawa coupling ($e^+ e^-$

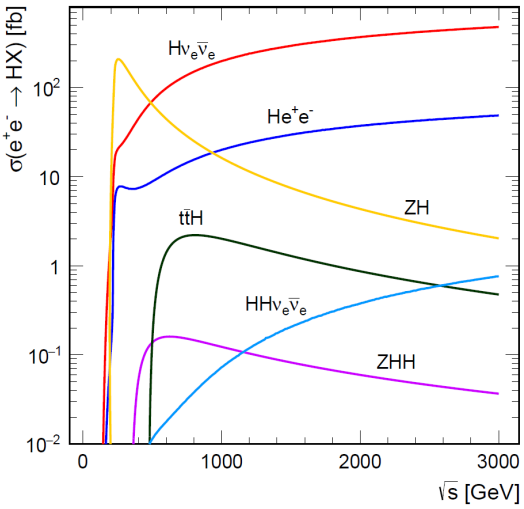


Figure 3.7: The cross sections of the main Higgs production processes at an e^+ e^- collider as a function of the centre-of-mass energy for a Higgs mass of 126 GeV [20].

$\rightarrow$ t $\bar{t}$ H) and the Higgs trilinear self-coupling (e^+ e^- $\rightarrow$ H H ν_e $\bar{\nu}_e$), for which the lowest order Feynman diagrams are shown in figure 3.9. The expected precision on the top-Yukawa is 4.2% for 1.5 ab^{-1} at $\sqrt{s}$=1.4 TeV [20]. This is expected to be improved to 4% with an electron polarisation of -80% [34]. As shown in figure 3.7, the cross-section decreases with energy, hence the precision is not expected to be better at the higher energy stage [20].

The measurement of the trilinear self-coupling gives access to the coupling λ assumed in the Higgs mechanism:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{3.1}$$

where μ is a parameter of the Higgs potential related to the mass of the Higgs boson. This is a crucial measurement to make when looking at the SM Higgs mechanism and is

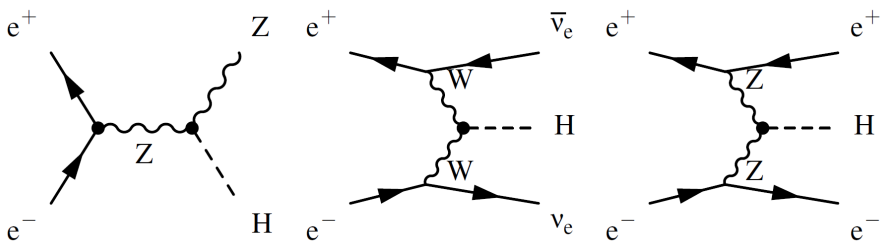


Figure 3.8: Feynman diagrams of the three highest cross section Higgs productions at CLIC: (a) Higgsstrahlung, (b) WW-fusion and (c) ZZ fusion [20].

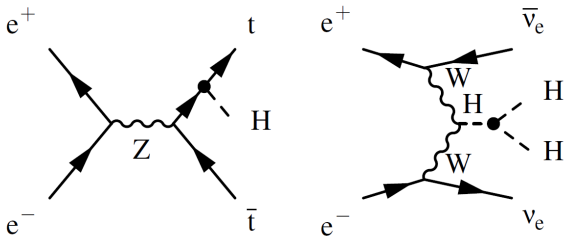


Figure 3.9: Feynman diagrams of the main processes involving: (a) top-Yukawa coupling g_{Htt} and (b) Higgs trilinear self-coupling λ [20].

particularly challenging at the LHC even with 3000 fb^{-1} of data [35]. It is also important as BSM physics models can introduce deviations of λ from its SM value of up to tens of percent [36]. The precision that can be achieved on the Higgs trilinear self-coupling, by combining both the 1.4 TeV and 3 TeV energy stages, is 26% for unpolarised beams and 19% for electrons with a polarisation of -80% [20].

To gauge the performance of CLIC, the coupling parameters of the Higgs have been extracted from several studies laid out in [20]. The fits were done only considering the statistical uncertainties of the measurements, not including systematic or theoretical ones, and performed in two ways: model-independent and model-dependent. The model-independent way means that the fit is performed with free parameters, relying on minimal assumptions from theory. Since model independent is unique to lepton colliders results from this method will be presented. Figure 3.10 shows the couplings for the three energy stages, with results from each subsequent stage building upon the previous one, allowing for a thorough examination of the newly discovered particle. With the aforementioned results achieving precision down to the percent level, CLIC is able to provide sensitivity to possible BSM phenomena that would cause deviations from the SM Higgs couplings, such as supersymmetry, mixed-in hidden sector Higgs bosons, and a composite Higgs boson [36, 37]. More details about the potential Higgs physics that can take place at CLIC are given in [20].

3.4.2 Top physics

Another important aspect of the CLIC programme is top physics. The top quark is the heaviest of the elementary particles in the SM with a mass of around 173 GeV and decays through the weak force before hadronisation occurs. The emphasis of the top quark studies will be the determination of the top quark mass to a very high precision, as well as the couplings. One of these couplings has already been mentioned, the top-Yukawa coupling, while others are the couplings to the photon and Z-boson. There is also a potential for more exotic decays to be studied such as $\mathrm{t} \rightarrow \mathrm{c}\,\mathrm{H}$, where CLIC may

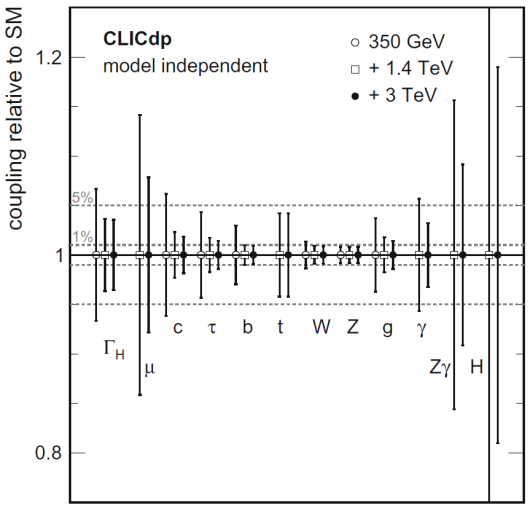


Figure 3.10: The precision of the Higgs coupling for the three energy stages obtained from model-independent fits without systematic and theoretical uncertainties [20].

provide competitive constraints to the full LHC programme, more details can be found in [38] and [39].

A key property of the top quark is its mass. One method to measure the mass of the top quark is to perform a top-pair production threshold scan which cannot be done at hadron colliders due to the parton-parton centre-of-mass energy not being known. This will be done in the first phase of CLIC, when a dedicated scan in steps of 1 GeV around $\sqrt{s}$=350 GeV will collect 10 fb^{-1} of data per step. As shown in figure 3.11(a), the $\mathrm{t\bar{t}}$ cross-section can then be studied as a function of centre-of-mass energy. The main advantage of this is that it provides a precise top mass measurement with well controlled theory uncertainties compared to top mass measurements above the threshold [22]. A precise measurement of the top mass is important because it impacts on theory predictions of production cross-sections necessary for both Higgs properties and BSM. The current LHC results yield an experimental uncertainty of 480 MeV [40] and the projection for the HL-LHC is 170 MeV [41], while for CLIC it is expected to be below 100 MeV [22]. The top mass can also be used as a test for the SM and to constrain any possible BSM, more details can be found in [42].

The top-quark couplings to the electroweak gauge bosons can be studied in detail by measurements of top-quark pair production above the threshold. The pairs are produced through s-channel processes via Z-boson and γ exchanges, a Feynman diagram for this process is shown in figure 3.11(b). The couplings to these bosons can therefore be obtained down to sub-percent precision [43], exceeding the precision expected from the LHC by an order of magnitude. Since the top-quark couplings to the electroweak

gauge bosons are tightly constrained in the SM, whilst in many BSM models there can be substantial deviations from these values, CLIC has a good discovery potential for new physics [23].

The branching ratios of exotic flavour changing neutral currents can also be measured at CLIC due to the clean environment found at lepton colliders. These types of decays are heavily suppressed in the SM model but are significantly enhanced in many BSM models, of the order of 10^{-2}–10^{-4} [44]. The LHC has a much more favourable rate of top quark pair production and lepton colliders will not be able to compete for decays where the final state produces distinct signals. However, for decays with less distinctive features, such as hadronic channels, CLIC could be superior [43]. For the $t \to c\,H$ decay the HL-LHC is expected to give a branching ratio limit of 2×10^{-4} [45], while at CLIC at 380 GeV with 500 fb^{-1}, the current reach is 1.4×10^{-4} [44].

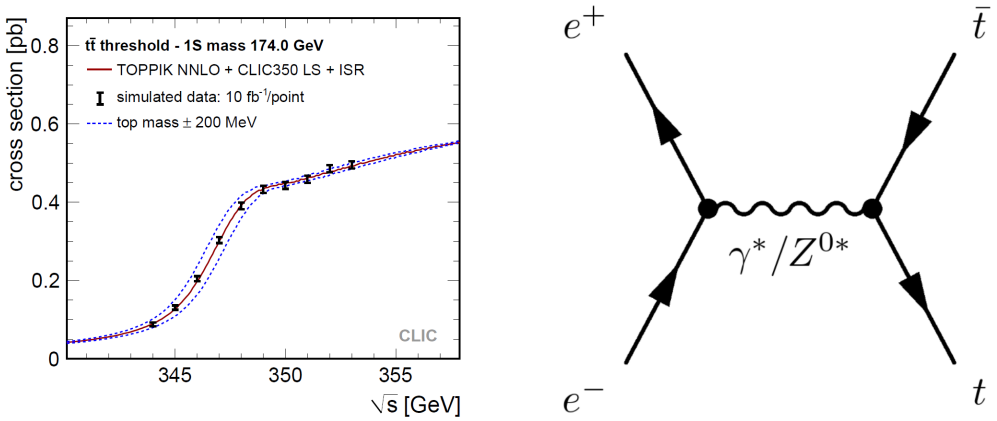
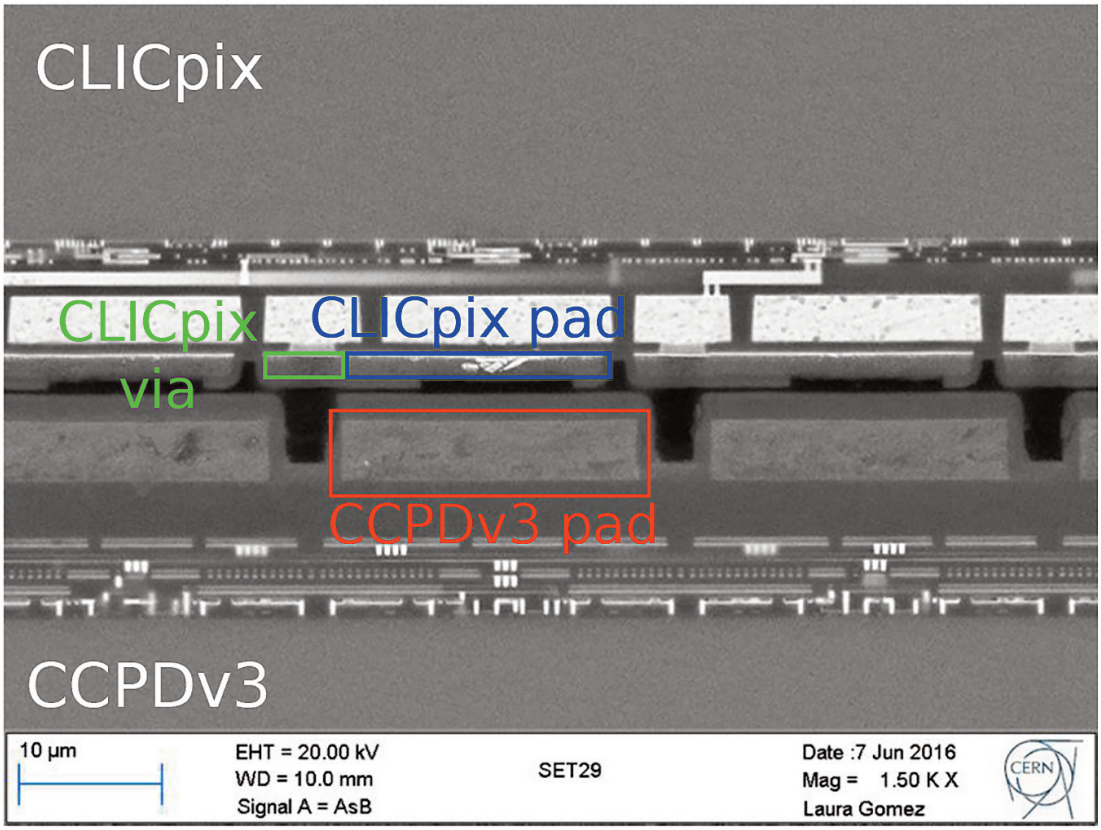


Figure 3.11: (a) The simulated $t\bar{t}$ cross-section as a function of centre-of-mass energy in a scan at the threshold with 1 GeV steps collecting 10 fb^{-1} each [22]. (b) A Feynman diagram of the $t\,\bar{t}$ production at CLIC [43].



A cross-section of a capacitively coupled assembly that was taken using scanning electron microscopy performed by the CERN EN-MME-MM group. (see page 57)

CHAPTER 4
CLIC vertex detector

The principle function of the vertex detector at CLIC is to reconstruct displaced vertices of heavy quarks and tau-leptons by efficiently tagging their decays. It will also provide information which, in conjunction with the tracking detector, will allow for precise and efficient track reconstruction, even for low-momentum particles. These high requirements arise from the physics goals and the beam structure of CLIC and are addressed by an R&D programme that is pushing the technology to its limits. Extensive studies of various parameters such as detector technology, layout and mechanics have been carried out and are ongoing to find the best possible solution that satisfies all the competing goals [6]. The exact experimental conditions and physics requirements for the CLIC vertex detector influence the choice of layout and sensor technology, all of which will be described in this chapter.

4.1 Physics requirements and experimental conditions

The goals of the vertex detector can be achieved through precise measurements of the impact parameter and charge of the tracks that emerge from the decay vertices. Full detector simulation studies have shown these the goals can be met with a constant term of $a \approx 5\,\mu\text{m}$ and a multiple-scattering term of $b \approx 15\,\mu\text{m}$, derived from the canonical parametrisation:

$$\sigma_{d_0} = \sqrt{a^2 + b^2 \cdot \text{GeV}^2/(p^2 \sin^3\theta)}\,, \tag{4.1}$$

where p is the momentum of the particle and θ is the polar angle with respect to the beam axis [6]. In addition to this, very good coverage, especially in the forward region, is needed down to $\theta \simeq 7°$. This allows the tagging of very forward electrons, which are used in many of the physics analyses for efficient background identification.

To reach the necessary impact parameter resolution, simulation studies have shown that a single point resolution of $\approx 3\,\mu\text{m}$ and a material budget of $0.2\%\,X_0$ for the beam pipe and per detection layer are required [6]. This low amount of material will minimise particle scattering and corresponds to an overall thickness of 200 µm for the sensor, readout, mechanical support, cooling and cabling. A consequence of the material budget is that 100 µm of this will be silicon. Hence, the sensors and readout chips will need to be very thin, 50 µm each. Another, is that the cooling will need to use forced airflow of dry gas, either air or nitrogen [46], to minimise material. To enable

air cooling a very low power consumption of below 50 mW cm^{-2} is needed, which is foreseen to be accomplished by power pulsed operation of the readout chip. This takes advantage of the beam structure by turning off most of the components during the 20 ms gaps between bunch trains and powering them back up a few µs before the next one [47].

The dominant source of background for the vertex detector will be the production of hadrons in $\gamma\,\gamma$ interactions and the creation of incoherent electron-positron pairs. An important consequence of this is in the very forward region where secondary effects, such as backscattering from the LumiCal and BeamCal into the vertex detector region, create the need for optimisation of the beam pipe and end caps. The incoherent $e^+\,e^-$ pairs and $\gamma\gamma \rightarrow$ hadrons backgrounds are expected to lead to an occupancy of around 3% per bunch train in the innermost layers [48]. The associated large combinatorics complicate the flavour tagging and tracking. To aid with this, a time binning of each hit with 10 ns is required to minimise the confusion with track matching and to separate background hits from those due to physics events from primary collisions. To aid the track finding, the proposed readout chip will need to be able to do simultaneous time-of-arrival (ToA) and energy measurements via time-over-threshold (ToT). A summary of the requirements is given in table 4.1.

Beam induced backgrounds are also the leading contribution to radiation damage in the silicon detectors. The total ionising dose is dominated by the incoherent electron-positron pairs. For the non-ionising energy loss, it is dominated by the $\gamma\,\gamma \rightarrow$ hadrons in the barrel region and both incoherent electron-positron pairs and $\gamma\gamma \rightarrow$ hadrons contribute equally in the forward region. This results in a maximum flux of about $4 \times 10^{10} n_{eq}$ cm^{-2} yr^{-1} and a total ionising dose of 200 Gy yr^{-1} including safety factors of up to five, this is $\sim 1 \times 10^4$ below LHC levels [48].

Table 4.1: Summary of the requirements for the vertex detector.

Requirement	**Value**
Impact parameter resolution	$a \approx 5$ µm, $b \approx 15$ µmGeV
Single hit resolution	≈ 3 µm
Material budget	0.2% X_0 per detection layer
Power consumption	< 50 mW cm^{-2}
Time binning	10 ns

4.2 Layout

The layout of the vertex detector in the CLICdet model is shown in figure 4.1. It is formed of three layers in a cylindrical barrel shape with three spiral disks on either side. In both the barrel and the disks there are two layers of active material per support structure hence, there are 6 detection layers in total for the barrel and disks. The reason for the spiral geometry is to allow air to flow from one end of the detector to the other for cooling [49]. Each disk consists of 8 segments, 24 in total for each side, that are arranged to create the spiral shape and which possess the same rotation direction at either end. As a result, there is an asymmetric layout of the disks. Detailed studies of the cooling have been carried out [46] and show that it is necessary for the airflow to be spiralling around the barrel layer to obtain sufficient cooling. Hence, the air must arrive to the vertex detector with a rotational component. This is achieved by spiralling air guides on the beam pipe.

The barrel is 260 mm long with a total area of 0.487 m^2 for the silicon sensors, which are mounted onto staves. To achieve the required flavour tagging performance, the innermost barrel layer needs to be placed as close as possible to the beam pipe while still being below the occupancy limit of the detector technology. The radius was chosen to be 31 mm. In the forward regions the disk segments (petals) are trapezoidal in shape and are at 90° to the detector axis. They are arranged such that there is an overlap of 2 mm from petal to petal with a longitudinal distance of 5.5 mm. The location of the first petal is 160 mm away from the interaction point and the last petal is located at 299 mm. For both sides combined, the total silicon sensor area is 0.351 m^2.

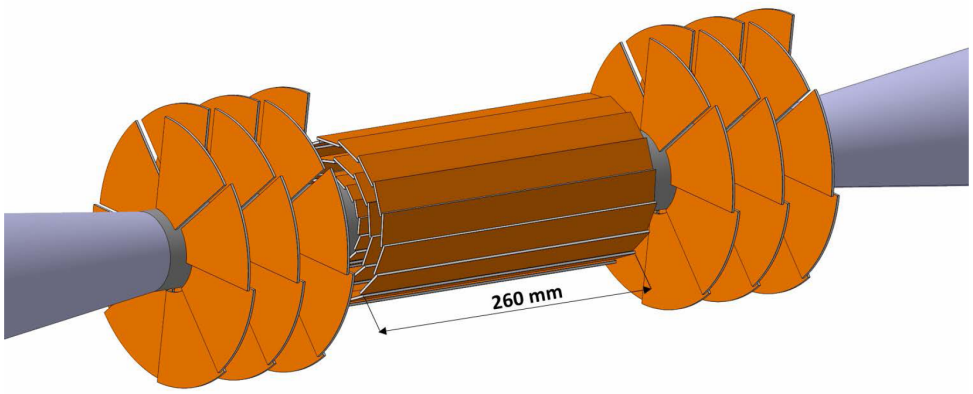


Figure 4.1: A schematic of the CLIC vertex detector showing the three double layers in the barrel and the spiral disks [28].

4.3 Proposed sensor technologies

In order to meet the demands of the vertex detector, an R&D programme has been set up to explore various technology options. While some of the demands have been met, a detector satisfying all of them is yet to be defined. The technological challenges for the sensors are: they must be very thin, 50 µm, and be able to collect the charge quickly to allow for fast signal processing and time binning of 10 ns. The main avenue of research pursues a hybrid sensor with a pixel pitch of 25 µm and digital readout. The choice of pitch is driven by the occupancy levels and by the single point resolution of 3 µm. The two hybrid concepts being developed are an interconnection between the sensor and readout chip using either small pitch bump-bonding to planar sensors or capacitive coupling to high-voltage CMOS (HV-CMOS) sensors.

4.3.1 Readout chip

The CLICpix ASIC is the first prototype for a hybrid readout chip matching the requirements of the CLIC vertex detector. Its design [50] is based on the Timepix chip [51]. It is implemented in a 65 nm CMOS process technology and consists of a 64×64 pixel array with a pitch of 25 µm and a thickness of 280 µm. In each pixel there is an analogue part which contains the CSA and the discriminator, along with a digital part that allows for the simultaneous measurement of a 4-bit ToA and 4-bit ToT, as shown in figure 4.2. The global threshold applied to all the pixels simultaneously is controlled by a 12-bit DAC on the periphery of the chip. In addition to this, there is also a local threshold set using a 4-bit DAC located in each pixel. The CLICpix also the option for data compression for each hit to reduce the readout bandwidth.

The operation of the chip is similar to a camera in that while the shutter is open it collects the data, then once the shutter is closed the chip reads out the data ready for the next open shutter. The ToT measurement is done by measuring the number of clock cycles during which the signal is above a certain discriminator threshold. Timing information is achieved by measuring the difference between the time the signal crosses the threshold and the shutter closing, effectively giving the ToA of the hit. In addition to this, there is also a counting mode on the chip where the counter is incremented each time the charge signal in the discriminator exceeds the threshold. The typical rise time of the pulses is ~ 30 ns with an equivalent input noise of $\sim 60\,e^-$ [50].

The CLICpix chip design has been tailored to suit the beam conditions found at CLIC, specifically the timing of the bunch trains. To reduce the power consumption of the chip, power pulsed operation has been implemented. Tests of this chip have been carried out and show that the performance is in accordance with the simulations [53].

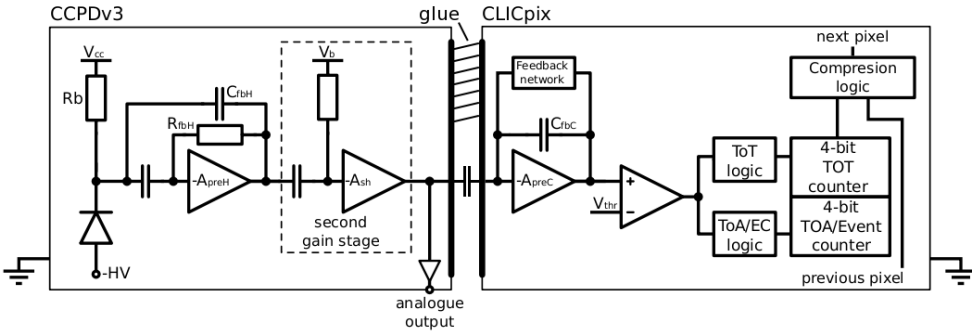
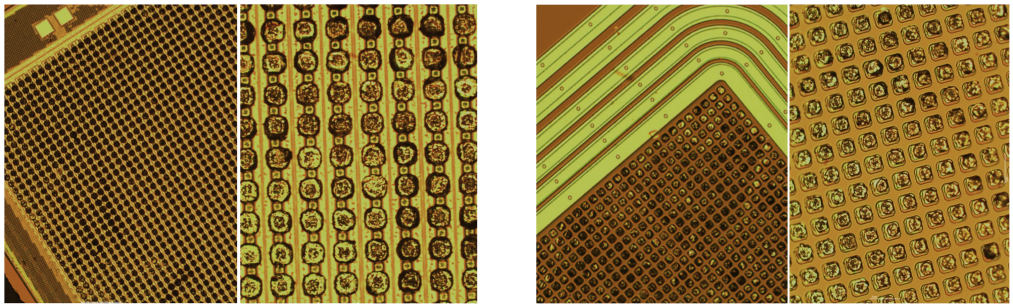


Figure 4.2: A schematic of the CCPDv3 and CLICpix circuitry found in one pixel [52].

4.3.2 Planar sensors

One option for the CLIC vertex detector is the use of planar sensors bump-bonded to the readout chip. Slim-edge sensors from Micron [54] and active-edge sensors from Advacam [55] have been produced with $55 \times 55\,\mu m^2$ pitch and thicknesses from 50 µm to 300 µm, in order to test the feasibility of very thin sensors using Timepix and Timepix3 [56] readout ASICs. The active-edge sensors attempt to reduce the size of the ineffective detection region at the edges and hence reduce the need for overlaps, reducing the material budget. Various edge layouts and guard ring structures have been tested, results of which can be found in [57].

A further set of sensors were produced using an ASIC specifically designed for CLIC, the CLICpix ASIC, bump-bonded to planar n-in-p pixel sensors. One set of sensors were produced by Micron [54] with a slim-edge design and a thickness of 200 µm. The other sensor was produced by Advacam [55] with a active-edge design with a thickness of 50 µm. The pitch of the sensors is 25 µm and they have a matrix of 64×64 pixels to match that of the CLICpix. The bonding process was done on a chip-to-chip level carried out at SLAC [58]. This type of fabrication, with such a fine bump pitch, is not industry standard, resulting in the need for new techniques to be developed and as a result, the assemblies received are considered as proof-of-principle. This process is achieved by first depositing the under-bump-metallisation (UBM) and indium bumps onto the pixel pads using a support wafer, then the two chips are connected using flip-chip reflow techniques. The two dies before the flip-chip process are shown in figure 4.3. These assemblies have been investigated and the results of the laboratory characterisations and calibrations are found in chapter 6.2.



(a) (b)

Figure 4.3: UBM and indium bumps deposited on (a) the readout chip (CLICpix) and (b) the slim-edge sensor prior to the flip-chip process [58].

4.3.3 Capacitively coupled pixel detectors

Connecting the sensor to the readout chip via bump-bonds comes with certain disadvantages, which become especially prominent for the sensor requirements and environments found at new colliders. One such disadvantage, is the extra processing steps and mechanical complexity that are involved in the bump-bonding process leading to a high cost. This is of importance since detectors are now using large areas of silicon for vertexing, CMS has an active silicon area of $4\,\mathrm{m}^2$ for the vertex detector [59]. Another disadvantage, is that the hybrid pixel size is limited by the current bump technologies, currently a limit of $20\,\mu\mathrm{m}$ for wafer-to-wafer bonding. To overcome these, a bumpless hybrid pixel detector that keeps the beneficial properties has been proposed, the capacitively coupled pixel detector (CCPD) [17, 60].

The concept of the CCPD is to bond the sensor to the readout chip using non-conductive glue and transfer the signal capacitively, a schematic of which is shown in figure 4.4. Once the signal is collected and amplified in the sensor, it is passed to the sensor coupling pad in the last metal layer. Similarly, the coupling pad of the readout chip is also in the last metal layer and matches the dimensions of the sensor coupling pad. The removal of the bump-bonding process means that CCPDs are cheaper and have lower mass. However, the capacitance between the two chips reduces the signal transfer efficiency, resulting in large losses if $C_{out} \ll C_d$, where C_{out} is the coupling capacitance between the sensor and the readout chip and C_d is the detector capacitance. Another disadvantage is that the alignment must be accurate in order to minimise unwanted coupling to neighbour pixels. Also, the thickness and planarity of the device must also be well controlled to ensure good capacitive coupling and a uniform detector response.

While there are different options to incorporate some form of in-sensor amplification, this thesis will focus on the HV-CMOS technology. High-voltage CMOS sensors are a p-type silicon substrate implanted with a deep n-well that itself contains further implants to make up the pixel electronics. The electronics are made from a combination of PMOS and NMOS transistors. PMOS transistors can be directly added to the deep n-well, while the NMOS transistors have to be implanted into a p-well, as shown in figure 4.4. Pixels have no parasitic wells and can be spaced close enough to each other to achieve a 100% fill factor; they also contain the electronics that provides the amplification using a charge sensitive amplifier (CSA). The choice of sensor technology for the CCPDs is one that has a CSA embedded in each pixel to ensure that the signal is large enough to be transferred to the readout chip [60].

The deep n-well acts as both a collection diode for the electrons, and as shielding for the low-voltage electronics from the substrate bias, allowing for a high bias to be applied in comparison to other CMOS technologies, $>$10 V. An advantage of this is that a relatively large depletion region (for this technology) will form, reducing the detector capacitance, and hence noise, as well as producing an increased signal. The main method of charge collection of this sensor is via drift, which happens rapidly due to the high electric field from the bias voltage; as a result the charge collection is fast (estimated to be $\sim$ 200 ps [62]). Furthermore, the relatively small active depth permits the sensor to be thinned down to 50 µm without significant charge loss.

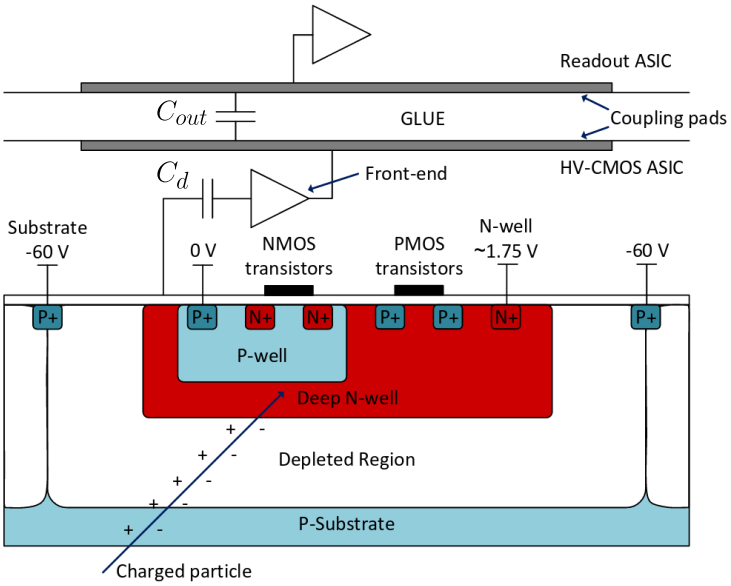


Figure 4.4: A schematic of the cross-section for a capacitively coupled pixel detector with a high-voltage CMOS sensor showing the deep n-well implant in the p-type substrate which hosts the electronics of the sensor [61].

The first prototype active sensor designed for CLIC was the CCPDv3 implemented in a commercial 180 nm high-voltage CMOS technology with a substrate resistivity of around 10 Ω cm and a thickness of 250 µm [62]. The analogue circuitry contained in each 25 ×25 µm^2 pixel makes up an integrating amplifier with a second gain stage. A schematic of the chip is shown in figure 4.2 along with that of the readout chip, the CLICpix. For the last four columns of the chip the second stage amplifier was not implemented, meaning that the voltage pulse was of a negative polarity and also reducing the gain by a factor of around two. The reason for this, was to test if such an architecture was feasible as this would greatly reduce the power consumption, something that is of importance for the CLIC vertex detector. No readout circuitry is incorporated in the sensor except for one row where a few select pixels (16 across the first row) allow the individual pixel analogue signal to be measured for testing. The whole matrix consists of 64 × 64 pixels. The foundry recommended maximal operating voltage of the device is –60 V, which is below the full depletion voltage.

4.3.4 Capacitively coupled assemblies

The CCPDv3 sensor and the CLICpix readout ASIC are connected to create a capacitively coupled assembly. This is done using flip chip techniques and the two chips are held in place by a thin glue layer of a few µm. This is done with the ACCµRA 100 flip-chip machine [63], allowing for precision alignment of 2 µm [52], and an epoxy resin [64] is used for the glue. Care is taken so that the two pads are aligned, as any misalignment could produce coupling to the wrong pixel and cause problems for the position finding algorithms. The two metal pads form capacitors which can transfer voltage signals by AC-coupling. A cross-section of a capacitively coupled assembly is shown in figure 4.5, providing a demonstration of the pad alignment. The total distance between the two pads is determined by the polyimide passivation layers as well as the glue, typically a few microns. The size of the CCPDv3 pad is 20 × 20 µm^2 while for the CLICpix the pad is 14 × 14 µm^2. On the CLICpix side there is also an additional part to the top metal layer, a via, that connects the input pad to the lower metal layers, shown in figure 7.1.

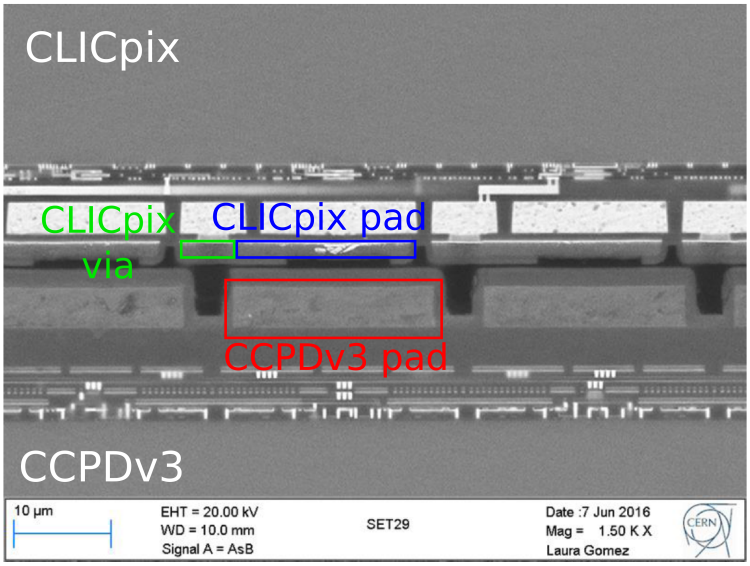
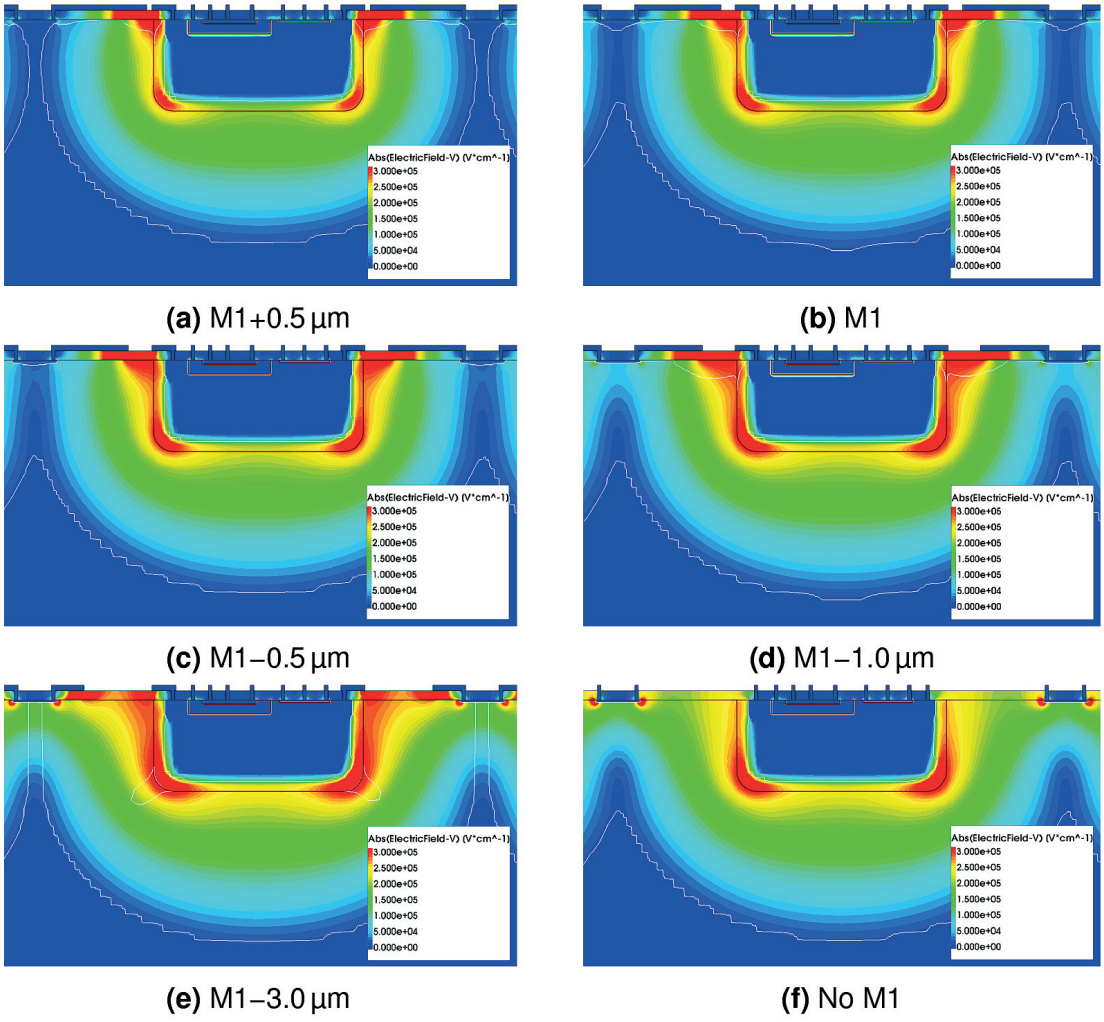


Figure 4.5: A cross-section of a capacitively coupled assembly that was taken using scanning electron microscopy performed by the CERN EN-MME-MM group [65].



(a) M1+0.5 µm **(b)** M1

(c) M1−0.5 µm **(d)** M1−1.0 µm

(e) M1−3.0 µm **(f)** No M1

The electric field of the simulations with different metal lengths of the sensor: (a) M1+0.5 µm, (b) no change (M1), (c) M1−0.5 µm, (d) M1−1.0 µm, (e) M1−3.0µm and (f) no first metal layer (No M1). (see page 72)

CHAPTER 5

TCAD simulations of HV-CMOS pixel structures

In this chapter a brief overview of Technology Computer Aided Design (TCAD) will be given. Details of the TCAD simulation method as well as results from the simulation of HV-CMOS pixel structures are presented. A 3-dimensional (3D) structure of a single-pixel is simulated in order to validate a simplified 2-dimensional (2D) single-pixel model. Other models simulated include: a three-pixel structure based on the validated 2D-model and finally a ten-pixel model. The ten-pixel model was used for angular simulations and comparison to beam test data which will be described in chapter 7. By performing TCAD simulations of the proposed sensors, a better understanding of the features of the measurements, such as transient signal development, are achieved. To reproduce the measurements an accurate model of the sensor was created in the simulation.

5.1 Overview of TCAD

In the semiconductor manufacturing industry, the design and production of devices is very complex, often having to be completed under very strict time and cost constraints. Development heavily relies on the number of engineering wafers or multi-project wafer submissions. In order to reduce this number, simulations are used to aid with the design. This produces a better comprehension of the process flow, the device operation and perhaps provide some insight into semiconductor behaviour, which can lead to optimised designs. It allows the electrical, thermal and mechanical properties of the devices to be predicted without the need for fabrication. One common tool to achieve this is TCAD [66].

TCAD is a powerful tool for simulating and optimising semiconductor fabrication processes and device operation. It does this by using fundamental physics models to produce the desired physical quantities of the device. To be able to model the structural properties and electrical behaviour of semiconductor devices, TCAD uses the finite element method. This technique first splits the structure into a grid, using a mesh. It then finds approximate solutions to the physical partial differential equations at each mesh point to calculate its physical properties, e.g. electric field and current. Typically, TCAD software consists of two main simulation tools: the `process` tool, used for the simulation of semiconductor fabrication processes and the `device` tool, used for the simulation of the electrical behaviour of the semiconductor. Often, the software also includes a Graphical User Interface (GUI), visualisation and plotting tools. There are

several simulation packages available which all offer common features. To integrate circuit elements into TCAD, SPICE (Simulation Program with Integrated Circuit Emphasis) [67] models are used and these types of simulations are called `mixed-mode`.

For this study Synopsys TCAD version I-2013.12 [68] was used. Both the 2D and 3D structures were generated using Sentaurus Structure Editor [69], which is a mixture of geometric and process emulation operations. From this the doping profiles and meshing strategies are defined ready for the next step of the simulation. After the structure is created, the electrical properties and device performance are simulated using Sentaurus Device [70]. This is a tool that is used to simulate the electrical, thermal and optical characteristics of silicon based semiconductor devices.

5.1.1 Process simulation

The process simulator is essentially a virtual semiconductor fabricator that models the effects of each step, such as: ion implantation, diffusion, growth and etching. It uses advanced models that have been calibrated with data to create a wide variety of structures that range from the nanoscale upwards [71]. Most process simulators are based on the industry standard `SUPREM-IV` code developed by the Integrated Circuit Laboratory of Stanford University [72]. The usual way to create a structure is to start by defining the boundaries for the chosen substrate and then, as in the fabrication process, go through each of the necessary process steps to produce the final structure. Another way is to use the geometric approach to define exactly the structure's features, e.g. implant profiles, aluminium layers etc., and where they should be. This has the advantage of being much quicker as it bypasses the lengthy diffusion process but can lack accuracy, especially when looking at the doping profile. In most cases, the geometric approach is used for large complicated structures where the exact fabrication process is not known, while the process approach is used for specific regions of the device or simple structures. The latter type of simulation can be used to check or calibrate the former. In the case of this study only the geometric approach was used.

5.1.2 Device simulation

The device part of the software can be used to simulate the electrical, thermal and optical characteristics of the semiconductor device [73]. It utilises the structure that was previously defined with the process simulator as an input and can be used to optimise a variety of semiconductor technologies, such as CMOS image sensors and heterojunction bipolar transistors (HBTs). It works by using the drift-diffusion model with possible additional models like: the hydrodynamic model (used for carrier energy transport), quantum mechanical models (used for quantum behaviour of devices such as quantum

tunnelling), or optical models (used for generation of light within the structure). The details of the various models are beyond the scope of this thesis but can be found in [70]. These various models can be added to the basic drift-diffusion model to enhance the application of the simulation to different scenarios. At each mesh point the simulation solves the Poisson equation:

$$\nabla \cdot \mathscr{E} = -\nabla^2 \varphi = \frac{\rho}{\varepsilon_r \varepsilon_0}, \tag{5.1}$$

where $\mathscr{E}$ is the Electric field vector, φ is the scalar electric potential field, ρ is the electric charge density, ε_r is the relative permittivity and ε_0 is the vacuum permittivity. It also solves, self-consistently[1] with the Poisson equation, the two continuity equations for electrons:

$$\frac{\partial n}{\partial t} = \frac{1}{q}\nabla \cdot \mathbf{J}_n - R_n + G_n, \tag{5.2}$$

and for holes:

$$\frac{\partial p}{\partial t} = \frac{1}{q}\nabla \cdot \mathbf{J}_p - R_p + G_p, \tag{5.3}$$

where $\mathbf{J}$ is the current density, R is the recombination rate, G is the generation rate, with the subscripts denoting electrons (n) and holes (p), and n and p are the electron and hole densities, respectively [7]. It does this by the discretisation of the structure into small areas, elements, relative to the mesh and solves the three differential equations at each element. For each mesh point these are solved self-consistently for the electrostatic-potential and the electron and hole concentrations. At each node the continuous functions (φ, n and p) are represented by vectors of function values and the differential operators are replaced with difference operators. Hence, the problem reduces to solving 3N non-linear algebraic equations, where N is the number of nodes. More information can be found in [74, 75]. Due to these equations being coupled non-linearly, the solution cannot be achieved in one step. Instead, an iterative method is used (typically the Newton iteration method) to provide an approximate solution. This process begins with an initial guess and continues until the convergence criteria have been reached, producing the physical quantities of the device. A simplified flow of the device simulation is shown in figure 5.1.

[1] This means that the solutions found satisfy all equations in the system simultaneously.

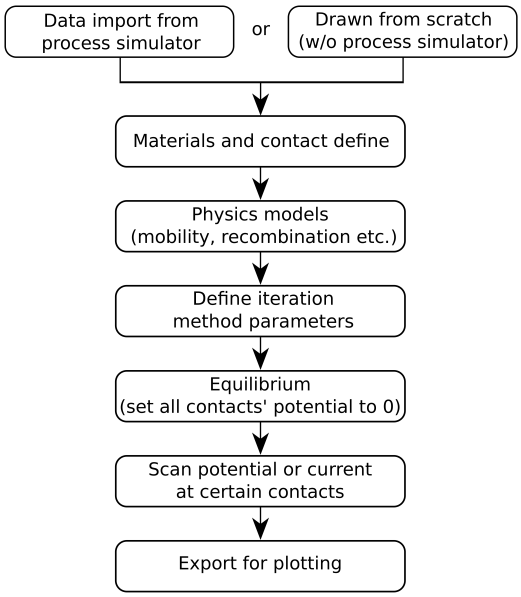


Figure 5.1: A simplified flow of the device simulation showing the basic steps involved [75].

The control file for the device simulation has four main parts: `Electrode`, `Physics`, `Math` and `Solve`. In the `Electrode` part, the contacts are defined along with their boundary conditions. The `Physics` part is used to define the physics models that will be implemented in the simulation, allowing for a parameterisation of the mobility, the recombination, avalanche, radiation damage and other models. The controls for the solver are defined in the `Math` section, where the numerical solver can be defined as well as other keywords specific to the physics models. Finally, the `Solve` section concerns setting the simulation completion criteria and solving for these using different processes.

The simplest process is to solve the Poisson equation or the Poisson and continuity equations for steady bias and time conditions. Another process is the quasistationary solver which ramps a parameter, typically bias voltage, in a series of steps until the defined goal is reached. For each step the solution is found for a steady state, i.e. constant time. The final process is a transient solver which simulates over time, often used for the signals produced by a particle traversing the sensor. As mentioned in chapter 5.1, a `mixed-mode` simulation can be performed. This is where extra devices and components (e.g. resistors and transistors) can be added and modelled by SPICE. One application of this is to use AC analysis to perform capacitance measurements [75].

5.1.3 Meshing

One of the key parts of the simulation, for both the `process` and `device` tools, is the mesh, as this determines the accuracy and efficiency [75]. A finer mesh means the simulation will be more accurate, but as a consequence there are: more calculations to be made, a longer simulation time and a larger memory usage. As a result, there is typically a trade-off between the accuracy of the model and the resources available hence, there is a need for an efficient mesh. Generally, the mesh will be finer around regions with a large gradient in the doping concentration and key areas for device operation, e.g. the channel region of a MOSFET, while a coarser mesh is found in the bulk. An example of a meshing strategy for an HV-CMOS pixel detector is shown in figure 5.2. There are many different ways to add to the basic rectangular mesh. Some of which are: refining the mesh according to the gradient of the doping concentration, refining at the boundaries between two materials and specifying user defined windows of finer mesh. The process of creating a mesh is to define the properties in the mesh file, once this is executed the mesh is automatically generated according to the criteria.

5.2 Comparison of 2D and 3D models

A comparison between a 2D model and a 3D model has been carried out to establish the validity of the 2D simulations. The reason for this is that the 3D simulations are very memory intensive, using large amounts of RAM ($>$ 64 GB) and as a result, take a long time to simulate ($>$ 1 week). Due to this, there is a trade-off between the mesh size, and hence convergence of the model, with the memory usage. Therefore, the 2D simulations are much quicker in comparison and are favoured for completing a large number of simulations. The 2D model simulates the width (x) and thickness (y) of the device while the 3D model simulates the width (x), length (z) and thickness (z), more details are given in the next subsection.

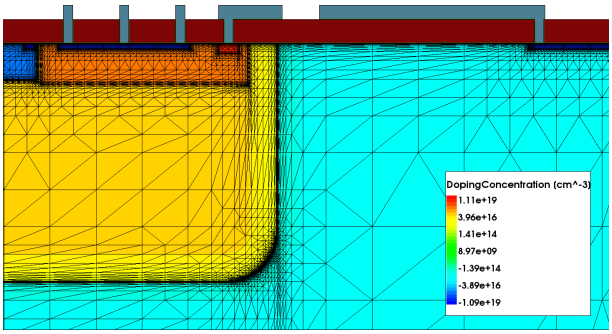


Figure 5.2: An example of the mesh in silicon for a HV-CMOS pixel detector. The mesh is refined according to the gradient of the doping concentration and also refined at the boundaries between the silicon, aluminium and silicon-oxide.

5.2.1 Simulated structures

The simulated structures were created by importing the graphic data system (GDS) layout file of the CCPDv3 into Ligament Layout Editor [76], which is able to open these types of files and extract the relevant information. These files are commonly used for chip design and are the ones that are submitted to the foundry. Within this tool, the unnecessary layers for the simulation, such as high-level metal lines and definition layers, were removed. This leaves only the implants, contacts and the first metal lines. A figure showing the resulting layers is given in figure 5.3. From this the main features of the sensor can be seen, the deep n-well surrounded by the p+ implant that makes up the bias ring. Ligament Flow Editor was then used to produce the masks for the implants, contacts and the metal that were utilised by Sentaurus Structure Editor to create the final structure. As the doping concentrations of the implants for the real device are not known, an estimate was employed, with the doping profiles being constant and having a Gaussian smearing towards the edge of the implant. Table 5.1 gives the parameters of the doping profiles used in the simulation. The p-bulk has a resistivity of 10 Ω cm.

In order to obtain a 2D model, a slice though the 3D implant structure needed to be defined. As the pixel is not symmetric, there are multiple options. A cut was chosen that goes through the main features of the pixel. In order to make the contacts, the position of the contact layer was adjusted, meaning that the cut chosen is not an exact cut of the CCPDv3. This is another reason why the 2D model needs to be verified by the 3D

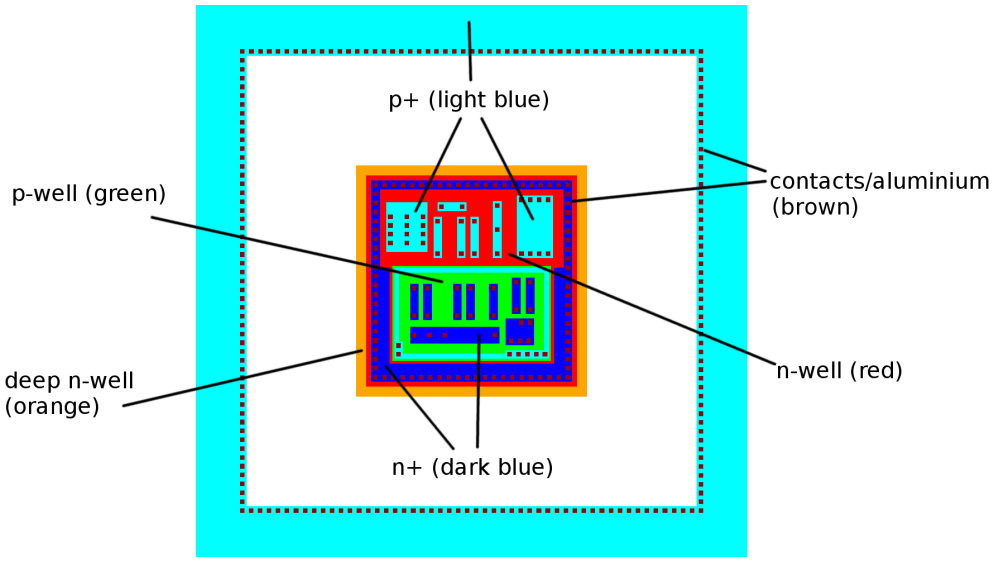


Figure 5.3: A top-down view, in Ligament Layout Editor, of the implant structure for one pixel of the CCPDv3. The deep n-well is in the centre and is surrounded by a p+ implant (bias ring) which is where the substrate bias is applied.

Table 5.1: The parameters of the simulated doping profiles.

Implant	Doping type	Concentration [cm^{-1}]	Gaussian width [µm]
Deep n-well	Phosphorus	1×10^{17}	1.6
n-well	Phosphorus	1×10^{18}	0.2
n+	Phosphorus	1×10^{19}	0.05
p-substrate	Boron	1.3571×10^{15}	N/A
p-well	Boron	1×10^{18}	0.2
p+	Boron	1×10^{19}	0.05

model. In this study three structures were simulated: full 2D, reduced 2D and reduced 3D, all shown in figure 5.4. The full 2D model is a cut of the 3D implant structure. The full 3D model was not simulated, for reasons discussed below. In its place a reduced model, which had some of the p+ and n+ implants removed (the implants that make up the NMOS and PMOS structures) was used. These implants are not expected to impact the results of the simulations in a significant way. To verify this assumption and to provide a cross-check, a reduced 2D model was also created that uses the same cut line as the full 2D model. This will help to identify if the difference between the full 2D and the reduced 3D is due to the removed implants or is due to the extra dimension of the reduced 3D model.

All the structures have a width (x) of 31.5 µm and a thickness (y) of 100 µm, with a length (z) of 31.5 µm for the reduced 3D model (for the 2D models this is effectively 1 µm). The thickness of the real device is 250 µm but to save memory a thinner structure was used. This is sufficient for this study as it is only comparing the 2D and 3D models. For comparisons to data the thickness has an impact on the timing of the charge collection as shown in appendix A, hence the real thickness was used. Another way to reduce the mesh in the 3D model was to create one continuous metal for the contacts, figure 5.4(c), as opposed to the series of squares shown in figure 5.3. The reason for these adjustments to the 3D model is that the amount of memory used was too large. This means the simulation would take a very long time or it would fail. Therefore, results for a 3D full model could not be obtained but the main features of the device are simulated in the reduced 3D model. The boundary conditions of the models are all the same. For the physical boundaries, the edges do not allow any charge to flow out. In the case of the electrical boundary conditions, the contacts are all held at a constant voltage, either being grounded or at a potential, as illustrated in figure 5.4(a).

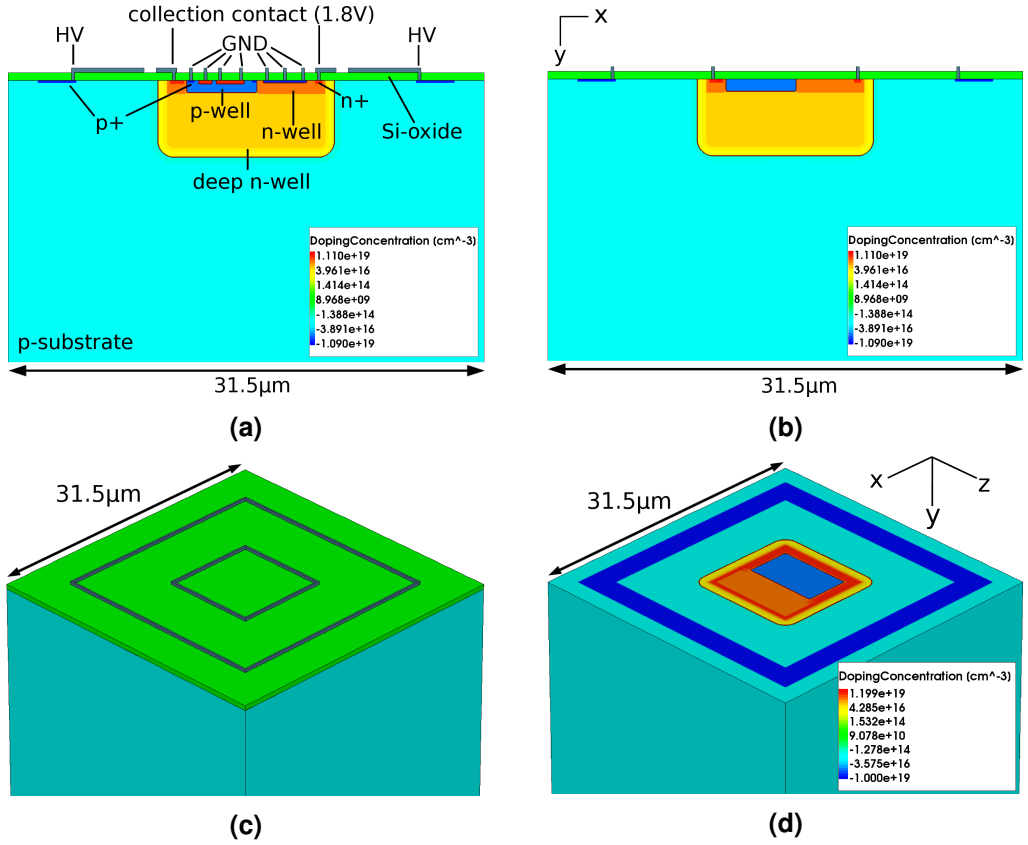
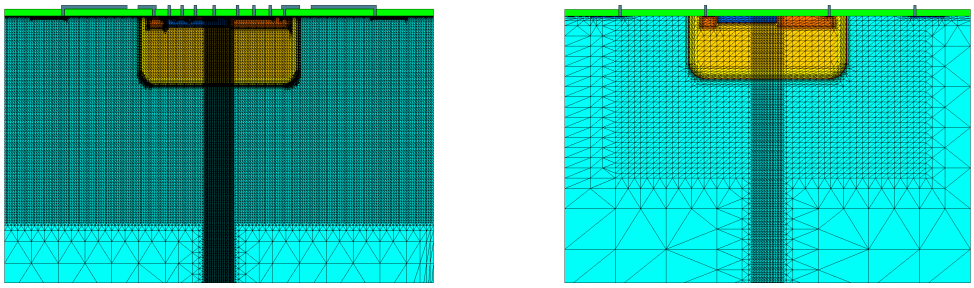


Figure 5.4: Four images of the three structures created using Sentaurus Structure Editor: (a) The full 2D model, (b) the reduced 2D model, (c) the reduced 3D model and (d) the reduced 3D model, with the metal and oxide removed for visualisation purposes.

5.2.2 Meshing

In addition to defining the geometry and doping profiles of the structure, a mesh is also defined. The mesh of the three models are all created using the same methods. However, in order to reduce the memory size in the 3D model, a larger mesh size was applied than those of the 2D models. A global mesh that refines around variations in the doping concentration within the silicon bulk is first defined. This helps to ensure that the doping profiles of the implants are adequately resolved, avoiding inaccurate results. The chosen refinement function is $f(x) = \text{arcsinh}(x)$ which checks whether the function value difference between two mesh points is above a certain value and then meshes accordingly.

To aid with the electrical simulations of the model, two refinement windows are placed in addition to the global mesh. One was set so that a finer mesh was present throughout the depletion region, and the other was placed along the path of the particle for the corresponding simulation created. An example of the final mesh for the full 2D is given



(a) (b)

Figure 5.5: An example of the meshing strategy for the full (a) 2D model and (b) reduced 3D.

in figure 5.5(a) and for the reduced 3D model in figure 5.5(b). In the 2D models the maximum mesh size was 5 µm and the minimum mesh size is 0.02 µm. Whereas for the reduced 3D model, the maximum mesh size was 5 µm and the minimum mesh size is 0.075 µm. In the generation of the 2D models' mesh, around 70k points are created, while for the reduced 3D model 145k point are used.

5.2.3 Electric field

To acquire the absolute value of the electric field, all three models were biased to –60 V (the foundry recommended maximal operating voltage of the device) using a quasi-stationary ramp. In all three structures the electric field showed the same trends, shown in figure 5.6. These are: high fields inside the depletion region and low fields outside the depletion region, as well as inside the deep n-well. The highest field region is found at the corners of the deep n-well. In the reduced 3D model, the high field also exists at the diagonal corners of the deep n-well implant near the surface, a feature that cannot be replicated in the 2D models. This is one of the advantages of the 3D simulation over the 2D simulation. The exception to this is that in the full 2D model, the high field extends up to and inside the oxide which can be attributed to the extended metal lines. The difference between the full 2D and the reduced 2D inside the central region deep n-well can be attributed to the extra implants found in the full model. However, in both cases the electric field in the central region of the deep n-well is negligible, $< 1 \times 10^{-5}\,\mathrm{V\,m^{-1}}$. The size of the mesh determines the resolution of the extracted values, a larger mesh means a smaller resolution. Due to this, the electric field map in the 2D cut of the reduced 3D model is not as smooth as the reduced 2D electric field map.

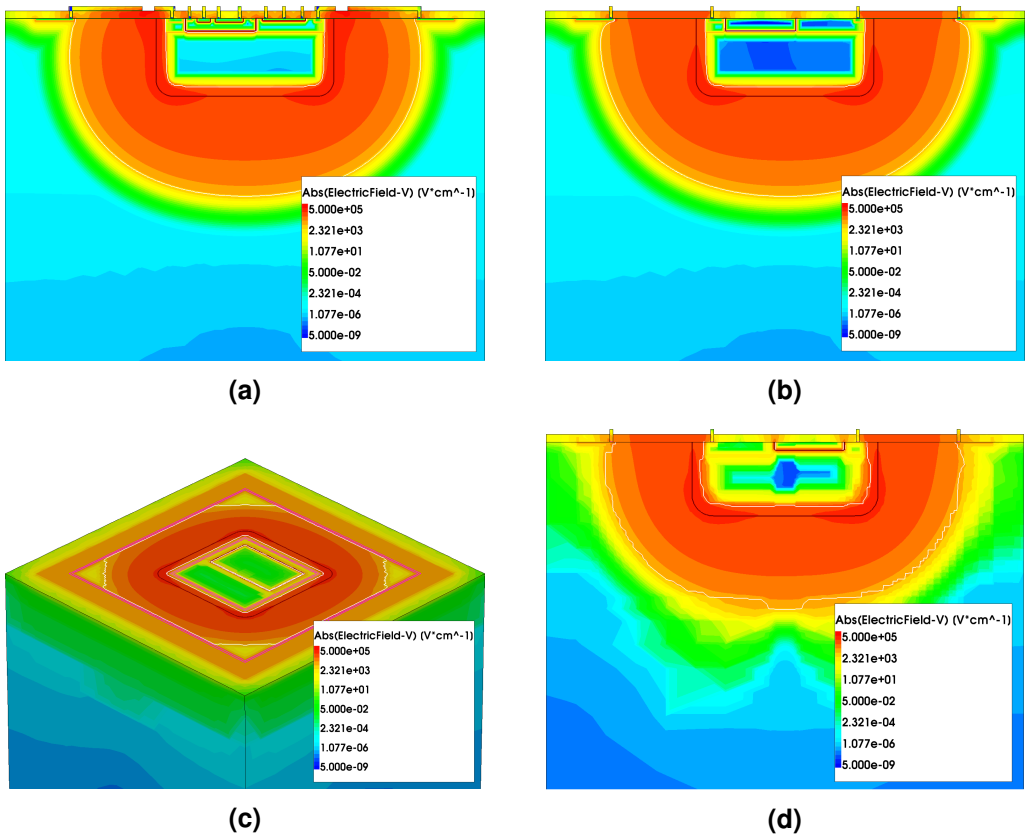


Figure 5.6: Absolute value of the electric field at an applied bias of –60 V: (a) the full 2D model, (b) the reduced 2D model, (c) the reduced 3D model and (d) a cut through the centre of the reduced 3D model.

5.2.4 Leakage current

Two important parameters of the sensor are leakage current and breakdown voltage, as they determine the bias voltage at which the sensor can be operated without detrimental effects on the performance, such as noise. Figure 5.7 shows the amount of current in the sensor as a function of applied bias voltage. To correct for the 2D nature of the 2D models, the current values were scaled by a factor equal to the length (z-direction) of the pixel, 25 µm. This is because the 2D simulations produce units of A µm^{-1} and the in-built scale factor (`AreaFactor`[1]) was set to 1. The current is very low compared to those in the measurements of typical sensors as the simulations are only for a one-pixel cell. They would need to be scaled according to the number of pixels on the sensor matrix to directly compare with data, as done in chapter 5.3.3. The breakdown voltage, corresponding to a large increase of current, is at $\approx$ –88 V for the full 2D model, at $\approx$ –90 V for the reduced 3D model and at $\approx$ –118 V for the reduced

[1] For 1D and 2D simulations it specifies the extension of the device in the remaining one or two dimensions when scaling simulated parameters such as current and charge.

2D model. The type of breakdown that occurs in the simulation is the avalanche breakdown described in chapter 2.1.5. In the real device the breakdown was measured to be −93 V [52]. Therefore, there is a good agreement of both the reduced 3D model and the full 2D model with the measurements.

Figure 5.8 shows the absolute value of the electric field for each of the structures at their breakdown voltage. In silicon breakdown occurs at $\approx 3 \times 10^5\,\mathrm{V\,cm^{-1}}$. In the full 2D model this electric field value is found in the corners of the deep n-well, as with all structures, as well as in the oxide and at the surface near the edge of the deep n-well. A possible cause of this is the fact that the metal lines above the oxide are so close together. This is not the case for the reduced 2D model where there is only a large area of high field at the corners of the deep n-well due to the absence of the metal lines. In addition, the depletion region (shown by the white line in figure 5.8, as defined by TCAD) becomes distorted due to Zener breakdown and produces a high current density in this region. For the reduced 3D model, the breakdown value of the field is seen at the corners of the deep n-well near the surface, which it cannot be replicated in the 2D simulations.

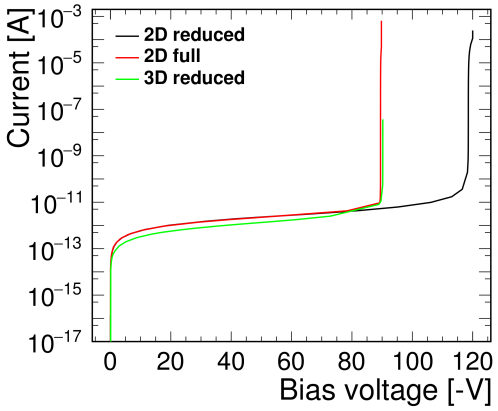


Figure 5.7: The leakage current as a function of the bias voltage for all three structures. Breakdown is indicated by the sharp rise in current.

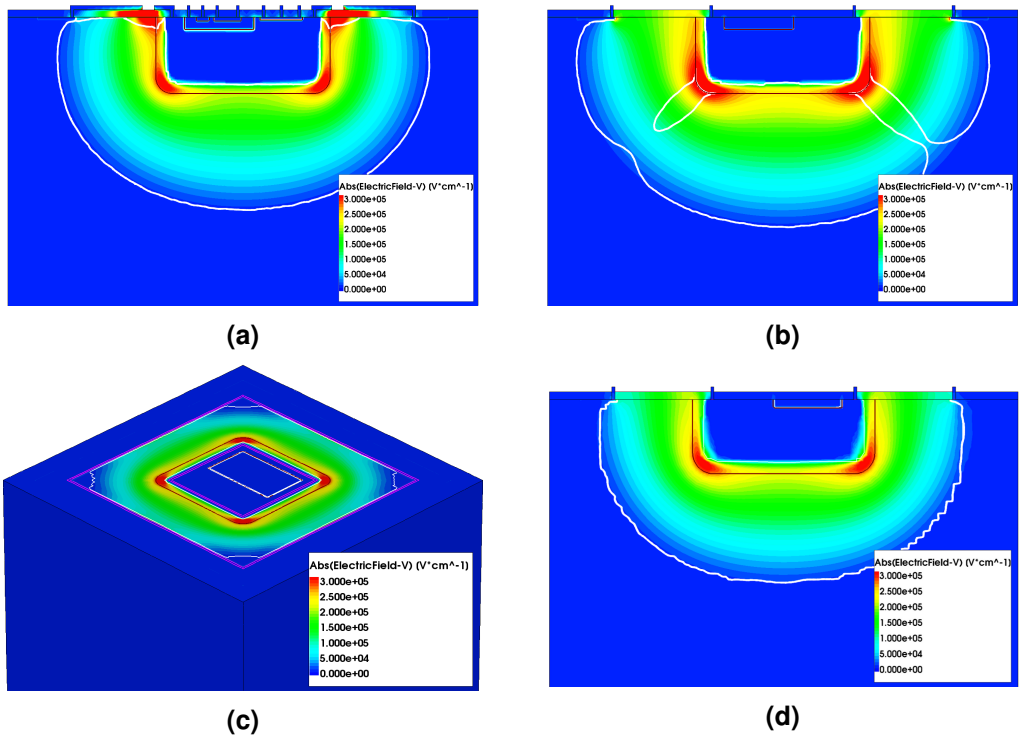
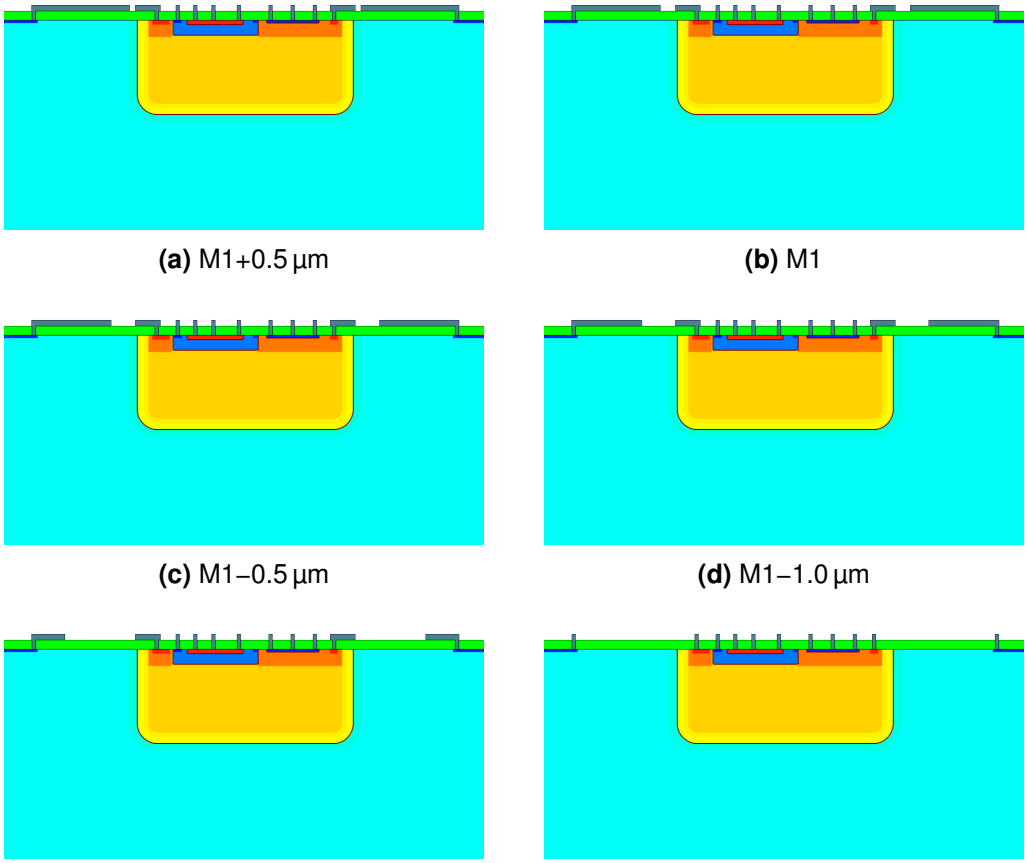


Figure 5.8: Absolute value of the electric field for all three structures at their breakdown voltage: (a) the full 2D model at −88 V, (b) reduced 2D model at −118 V, (c) reduced 3D model at −90 V and (d) a cut though the centre of the reduced 3D model at −90 V. The white line represents the edge of the depletion region.

Effect of metal line distance

The above comparison of the current-voltage characteristics shows the difference between the breakdown of the full 2D and reduced 2D models. One feature that was estimated to contribute to this difference was the length of the metal lines that sit on top of the oxide and are part of the first metal layer (M1) in the GDS file of the sensor chip. To investigate this further, a series of simulations with different metal lengths were produced, as seen in figure 5.9, in order to substantiate this claim. Six models were created in total: M1+0.5 µm, no change (M1), M1−0.5 µm, M1−1.0 µm, M1−3.0 µm and no first metal layer (No M1). The numbers after M1 denote the change in length of the M1 lines in comparison to the nominal value.



(a) M1+0.5 µm **(b)** M1

(c) M1−0.5 µm **(d)** M1−1.0 µm

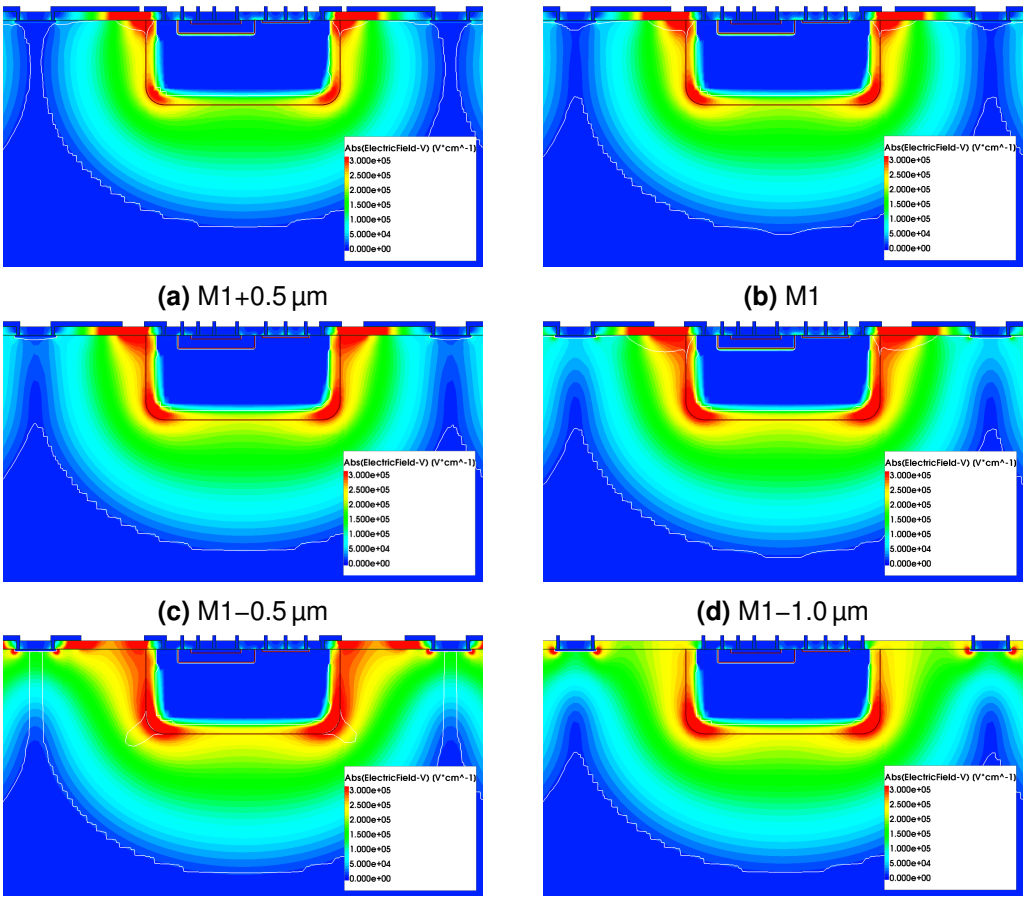
(e) M1−3.0 µm **(f)** No M1

Figure 5.9: The various length of metal lines in the first metal layer of the sensor: (a) M1+0.5 µm, (b) no change (M1), (c) M1−0.5 µm, (d) M1−1.0 µm, (e) M1−3.0 µm and (f) no first metal layer (No M1).

A comparison of the electric fields of the various models at their breakdown voltage is shown in figure 5.10. The larger the distance between the metal lines the smaller the high field region found in the oxide and at the Si-oxide boundary, eventually disappearing completely when there are no metal lines. As a result, the breakdown voltage increases the larger the distance between the metal lines, as shown in figure 5.11. The difference between model M1 and model M1−3.0 µm is an increase in 50 V between the breakdown voltages. The curves for M1−0.5 µm and No M1 stop early because the simulation does not converge and fails. Contrary to the trend, M1−3.0 µm has a larger breakdown voltage than No M1 by around −10 V.

The reason for this is not entirely known and could be due to a few factors. The first reason could be due to the difference in the electric field but the No M1 model has a smaller high value of electric field region. Another could be an unphysical effect in the

simulation. A better indication comes from the hole density shown in figure 5.12. In the M1−3.0 µm model the hole current spreads directly downwards from the HV contact and punches through to the corner of the deep n-well. Whereas in the No M1 model this does not occur instead the breakdown must come from avalanche, as is the case with M1−0.5 µm. The reason for this is that the No M1 model and the M1−0.5 µm model do not display the same high value of the hole current at the surface seen in the M1+0.5 µm, M1 and M1−1.0 µm. In these cases this is the source of the breakdown, a short from the HV contact to the deep n-well, as seen by the large hole current density and distortion of the depletion region found at the surface between the HV contact and the deep n-well, shown in figure 5.12.



(a) M1+0.5 µm **(b)** M1

(c) M1−0.5 µm **(d)** M1−1.0 µm

(e) M1−3.0 µm **(f)** No M1

Figure 5.10: The electric field of the simulations with different metal lengths of the sensor: (a) M1+0.5 µm, (b) no change (M1), (c) M1−0.5 µm, (d) M1−1.0 µm, (e) M1−3.0 µm and (f) no first metal layer (No M1).

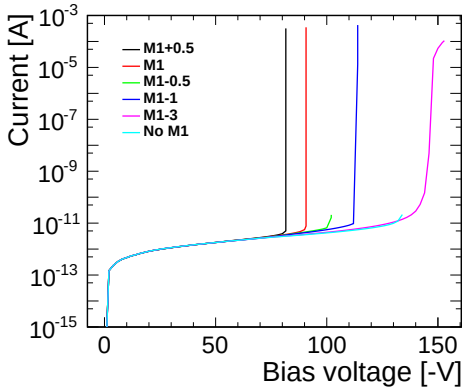
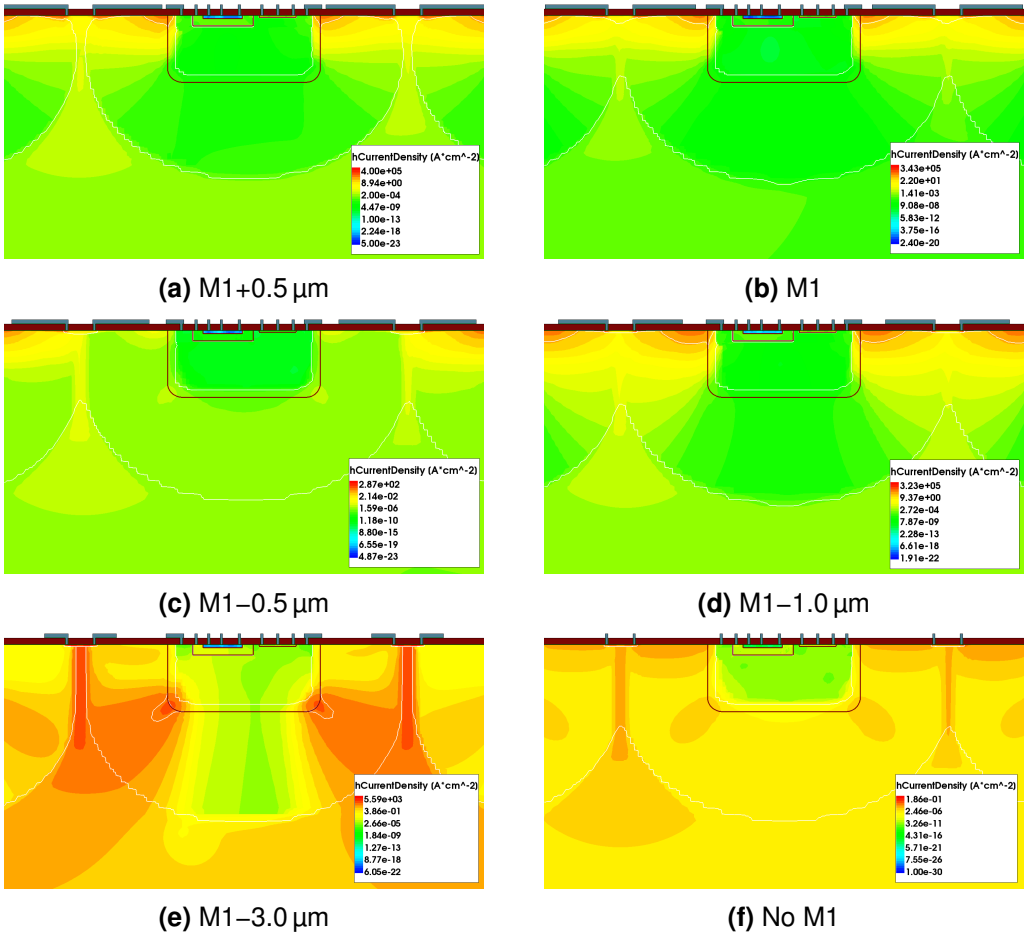


Figure 5.11: The current as a function of voltage for a the various metal line lengths.



(a) M1+0.5 µm **(b)** M1

(c) M1−0.5 µm **(d)** M1−1.0 µm

(e) M1−3.0 µm **(f)** No M1

Figure 5.12: The hole current density of the simulations with different metal lengths of the sensor: (a) M1+0.5 µm, (b) no change (M1), (c) M1−0.5 µm, (d) M1−1.0 µm, (e) M1−3.0 µm and (f) no first metal layer (No M1).

5.2.5 Capacitance

The deep n-well to bulk capacitance value was simulated using small signal AC analysis and the mixed mode of Sentaurus Device at a frequency of 1 MHz. It measures the magnitude and phase of the currents in the circuit in response to a small periodic voltage excitation. Calculations from geometric considerations estimated this value to be $\sim$10 fF at –60 V. The results of the simulation shown in figure 5.13 are consistent with the calculations. At –60 V the 2D models have similar values at around 5 fF while the reduced 3D model has a higher value of 8 fF. The difference between the 3D and 2D models is around 40% and with the 3D model being closer to the estimated value. Again, this further shows the advantage of the 3D simulations over the 2D simulations. The reduced 3D model's response is jagged due to the coarser mesh. The values obtained from the 2D models are obtained in the units of F µm^{-1} and are then multiplied by the length of the deep n-well to produce the final value and account for the 2D nature of the simulations. However, the deep n-well sidewall to substrate contribution in the z-direction also needs to be accounted for, which is not possible in the 2D models. For these reasons they can only be considered as partial estimates.

5.2.6 Response to a Minimum Ionising Particle

To simulate a Minimum Ionising Particle (MIP) in TCAD the `Heavy Ion Model` of Sentaurus Device is used with the time, direction, position and charge deposition of the particle specified in the command file. The temporal and spatial deposition of charge are modelled with Gaussian distributions. The sigma of the temporal Gaussian is 2 ps. Due to this, the time at which the MIP enters the transient simulation is not 0 s but 10 ps so that the signal does not become distorted and the current is allowed to settle. The

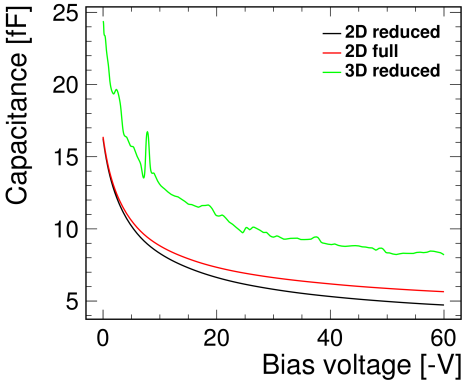


Figure 5.13: Capacitance as a function of the bias voltage at a frequency of 1 MHz for the three simulated models. The jagged line of the reduced 3D model can be attributed to the coarser mesh.

sigma of the spatial distribution is 0.035 µm. In this study the MIP passes through the centre of all three structures and deposits 80 electron-hole pairs per µm over a depth of 100 µm, 8000 e^- in total, with no Landau fluctuations taken into account. This value is the most probable value for energy loss in silicon, which is the most accurate quantity to use as discussed in chapter 2.1.6, and will be used throughout this thesis.

Other important parameters for the MIP simulation are the mobility and recombination models used as they impact heavily of the charge transport behaviour of the carriers. In this study the models used are given in table 5.2 [70]. This results in a lifetime of the minority carriers, electrons, taking a value of $\sim 8 \times 10^{-6}$ s in the bulk.

A transient simulation from 0–10 µs was performed at a bias voltage of –60 V, and the current obtained at the collection contacts was measured (each contact is Ohmic and has a resistance value of 1 mΩ). The current pulse shape in the first 0.5 ns is shown in figure 5.14. The reduced 3D model has the largest and fastest peak but quickly drops to the lowest value. At all times the full 2D model has a larger current value than the reduced 2D model, which is due to the metal lines and extra implants.

The collected charge is found by integrating the current over time, see figure 5.15. Initially the reduced 3D model collects the most. This is due to the coarse meshing in the 3D model not generating as smooth an electric field or Gaussian distribution of deposited charge. After 100 ns, the difference between all three models is around 300 e^- ($\approx$ 15%), with the reduced 3D model collecting the least and full 2D collecting the most. The reason for using 100 ns is that simulations have shown that the rise time on the CCPDv3 chip is approximately 100 ns [77] and the charge collected at this time is what contributes to the signal. After 10 µs the difference between all three models reduces slightly to around 200 e^-, which amounts to a difference of $\approx$ 3%. Also, after 10 µs, the reduced 2D model collects the least amount of charge; this change in order occurs at around 4 µs.

Based on all the result of the 2D-3D comparison, the full 2D model is a reasonable approximation of the 3D model and will be used for the three-pixel structure described in chapter 5.3 and also be used for the angular simulation in chapter 5.5.6.

Table 5.2: The mobility and recombination models used in the TCAD simulations [70].

Mobility	Recombination
DopingDependence	SR(DopingDependence, Tunneling(Hurkx))
eHighFieldSaturation	Auger(withGeneration)
hHighFieldSaturation	Avalanche(UniBo, Eparallel)
Enormal	Band2Band (Hurkx)
CarrierCarrierScattering	

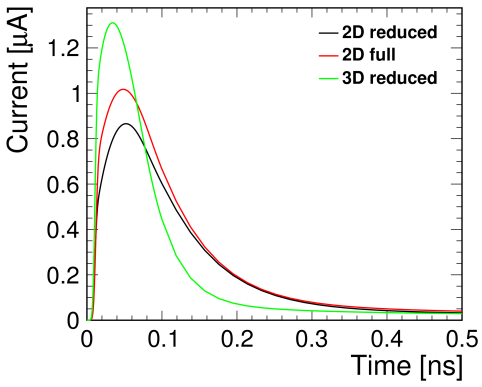


Figure 5.14: The current induced on the collection contact when a MIP passes through the centre of the pixel cell.

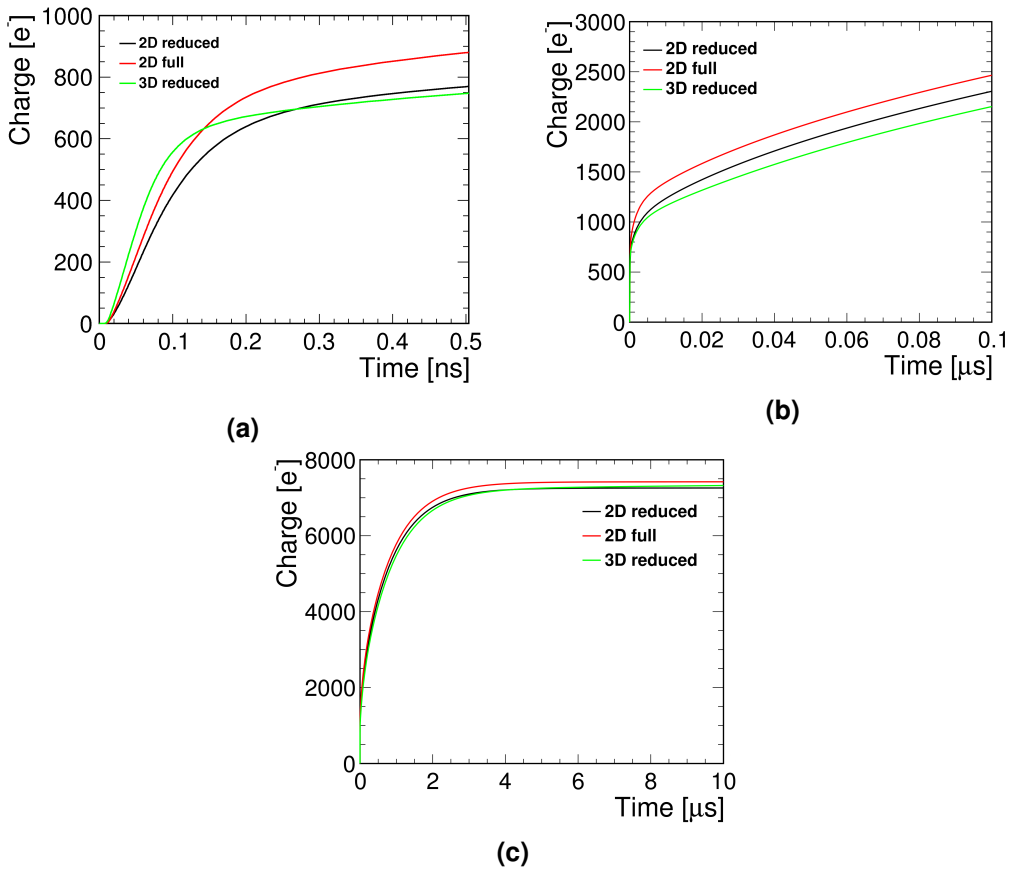


Figure 5.15: The total charge collected for the 2D-3D comparison at different times: (a) 0.5 ns, (b) 100 ns and (c) 10 µs.

5.3 Three-pixel structure

In this section the simulation method and results of the full 2D model are extended to include a three-pixel structure, with a bulk resistivity of 10 Ω cm. The motivation for producing a three-pixel structure is to study the charge-sharing between neighbouring pixels and to investigate the variation of the collected charge for different MIP positions across the pixel cell.

5.3.1 Simulated structure

The model was derived from the GDS layout file of the CCPDv3 sensor and created as described in chapter 5.2.1. The whole structure is 75 µm wide and 100 µm thick; the pixels have a pitch of 25 µm. The pixels are labelled 1, 2 and 3 from left to right, as shown in figure 5.16. In order to replicate a multi-pixel structure, periodic boundary conditions (PBCs) were added to the edges of the model. In this study the Mortar PBC approach (from Sentaurus Device) was used, this assumes both sides of the PBC interface (in this case the vertical edges of the model) are connected together. The simulation is (weakly) continuous across this interface, and currents can flow from one side to the other, more details can be found in [70]. This is an unphysical effect but better recreates charge flow in the pixel cell than without the PBC by preventing large charge build up at the edges. Although, as a result charge can flow from one side to the opposite side and produce a signal on the furthest pixel for large integration times, which is unphysical.

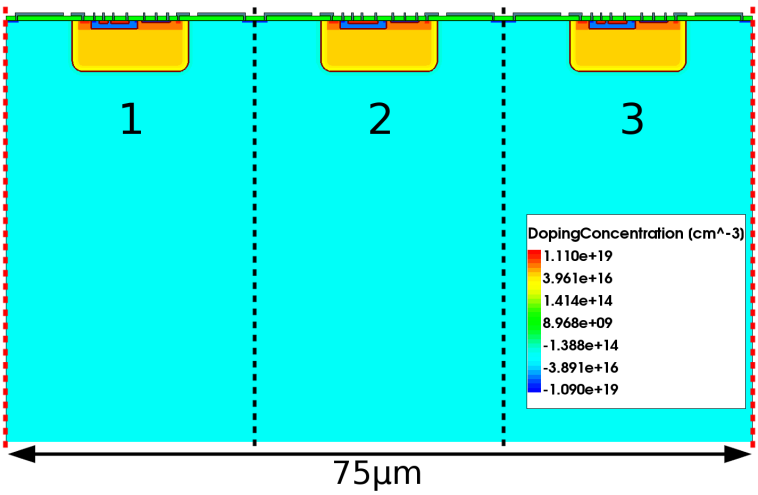


Figure 5.16: The three-pixel extension of the full 2D model with a pixel pitch of 25 µm and a bulk resistivity of 10 Ω cm. The pixels are labelled 1, 2 and 3 from left to right with the black dashed line indicating the boundary between them. The red dashed line represents the position of the periodic boundary conditions.

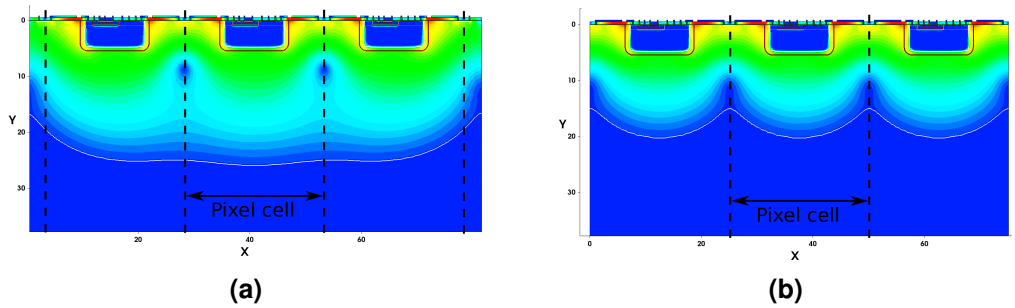
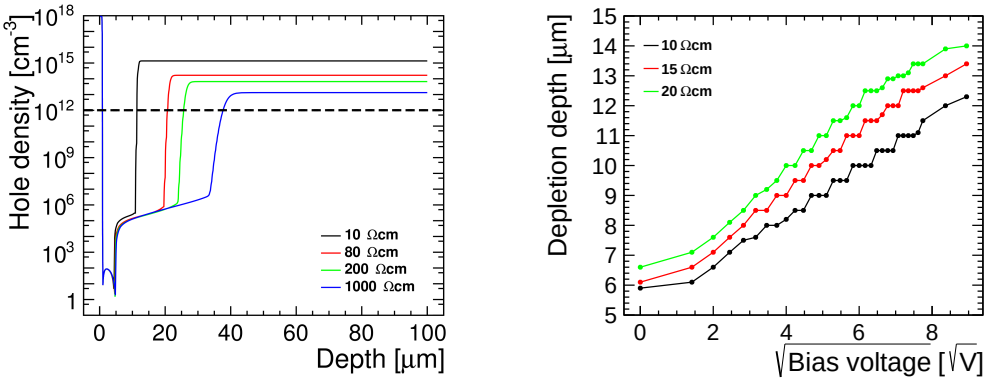


Figure 5.17: The electric field for a simulation (a) without PBC and (b) with PBC. The white line represents the depletion region.

As well as allowing current to flow through the pixel cells more realistically the PBC produce a more realistic electric field as it does not distort the field due to being close to a physical edge. This means that the electric field and depletion region are more symmetric within the pixel cell, as illustrated in figure 5.17. In addition to this, the depletion depth is not as large. In the case for no PBC conditions, the distortion of the depletion region can therefore be attributed to the edge effects of the simulation and would not be physical.

5.3.2 Depletion depth

Due to the sensors not operating at full depletion the exact depletion depth is an important quantity to simulate and measure as it is related to the charge collection and timing performance. The depletion width of a p-n junction is proportional to the resistivity of the substrate and the square root of the applied bias voltage, equation 2.29. The depletion depths of the simulated structures were obtained by extracting the hole density along a line going through the centre of the structures. If this value was above a threshold of $1 \times 10^{12}\text{cm}^{-3}$, then the edge of the depletion region was deemed to have been reached. An example of the hole density for the four bulk resistivities is shown in figure 5.18(a). This was done at various bias voltages to extract the depletion depth. Figure 5.18(b) shows the simulated values of the depletion depth obtained from TCAD. Edge transient current technique based on a two photon absorption process (TPA-eTCT) measurements of this value at –80 V give a depth of $(10\pm1)\,\mu\text{m}$ [78] compared to the simulated value of 12.3 µm, showing reasonable agreement between the two. The measured value of the resistivity of the device is $\sim 15\,\Omega\,\text{cm}$, being larger than the nominal value in the simulation, $10\,\Omega\,\text{cm}$. Therefore, figure 5.18(b) also shows simulations at $15\,\Omega\,\text{cm}$ and $20\,\Omega\,\text{cm}$. This difference between the measured and simulated values could arise from the limitations of the simulation such as it only being in 2D and the simulation being an idealised case not exactly replicating the doping profiles, which will have a large impact.



(a) **(b)**

Figure 5.18: (a) The hole density through the centre of the structures for different bulk resistivity values at −60 V. The dashed line is the threshold at which the edge of the depletion region is estimated. (b) The depletion depth as a function of the square root of bias for three bulk resistivity values.

5.3.3 Leakage current

The current-voltage characteristics are presented in figure 5.19, showing a comparison between the simulation and the data obtained from lab measurements for one device. Since the simulation was of a 2D one-pixel cell, the current was multiplied by 25, the length of the pixel. This corrects for the 2D nature of the simulations. The simulated value was also multiplied by 4096, the number of pixels on the device under test (DUT), to match the data. The current for the simulation and data agree very well, with the largest discrepancy being 5% before breakdown. The breakdown from the simulation is sharper than the data and occurs at −88 V, while for the data it occurs at −93 V resulting in a good agreement between the two.

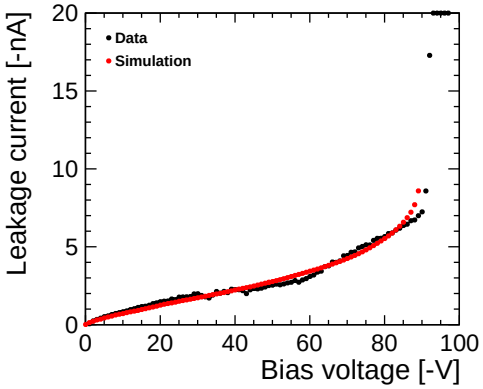
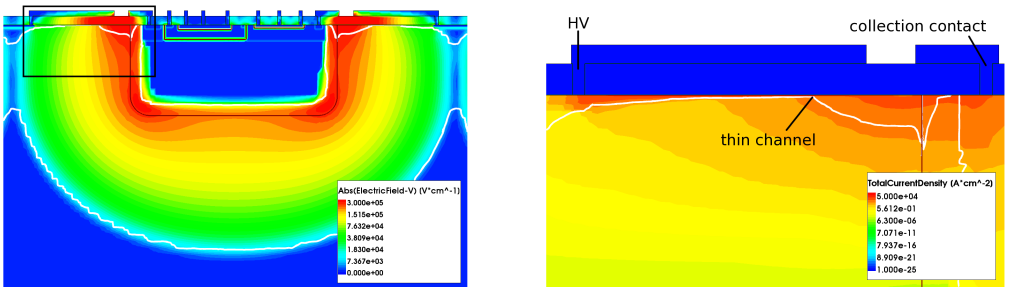


Figure 5.19: A comparison of the leakage current as a function of bias voltage between the TCAD simulation and data with breakdown at −88 V and −93 V respectively.



(a) **(b)**

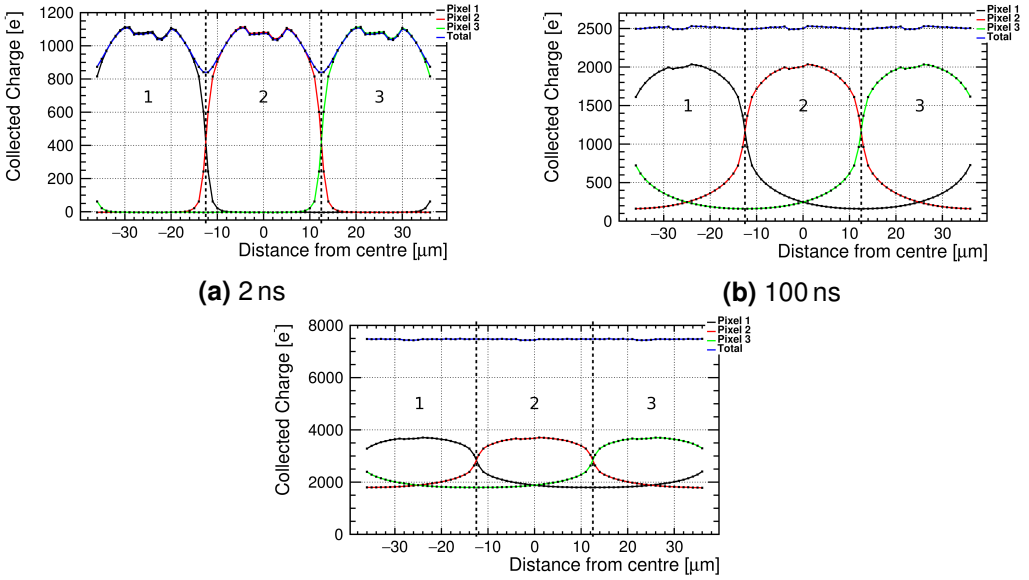
Figure 5.20: Two field maps at the breakdown voltage of –88 V showing (a) the absolute value of the electric field and (b) the total current density in the region near the high-voltage contact and the edge of the deep n-well, represented by the rectangle on (a).

In the simulations, at the breakdown voltage of –88 V, the high field regions already visible at –60 V become much more pronounced and extend further around the deep n-well, shown in figure 5.20(a). In the high field region near the silicon-oxide boundary the depletion region has become distorted, producing a thin channel through which current can flow. This allows a short between the high-voltage contact and the collection contact and is the reason for breakdown in the TCAD simulations. This is highlighted by the large value of the total current density in this region shown in figure 5.20(b). This information was used and has been optimised for the next generation of chips, the C3PD.

5.3.4 Response to a Minimum Ionising Particle

A series of MIPs entering perpendicular to the surface of the three-pixel structure were simulated and the charge collected after different times is shown in figure 5.21. The MIP entry positions were simulated from the far left at –36 µm to the far right at +36 µm, with the centre of the central pixel at 0 µm.

After 2 ns, all three-pixels collect the same amount of charge when a MIP passes through them, see figure 5.21(a). When the MIP passes next to the deep n-well, this position produces the largest charge collection. The slight increase at the edges of the deep n-well is because the electric field is the largest in this region compared to the rest of the pixel cell. In addition, the pixels collect zero charge when the MIP is sufficiently far away, indicating that no diffusion occurs from these regions yet. The lowest charge collection is seen under the bias ring (p+ implant). This is due to the depletion region not fully reaching these areas and a low value of the electric field. After 100 ns, figure 5.21(b), charge sharing starts to occur as diffusion becomes dominant. Furthermore, the two side pixels collect some charge when the MIP is simulated at the opposite



(a) 2 ns **(b)** 100 ns

(c) 10 µs

Figure 5.21: The charge collected by each pixel and the sum of all three (total) for the three-pixel model, for different MIP positions along the surface of the structure at different integration times : (a) 2 ns, (b) 100 ns and (c) 10 µs. The dashed lines represent the pixel boundaries.

edge, which is an unphysical effect. This effect is due to the PBCs which allow charge to flow through one side of the model to the other and is an artefact of the simulation. The total charge collected across the whole of the device after 10 µs is uniform within 1%. Charge sharing within the sensor helps with position reconstruction and has an impact on the performance of the final detector. Hence, it is important to characterise and quantify.

5.4 Angled simulations

Angular studies are needed to determine the performance of the HV-CMOS assemblies in the geometry of the CLIC vertex detector from a perpendicular incidence, up to a 80° incidence angle with respect to the perpendicular. For this purpose, the TCAD model was extended to include ten pixels to better replicate the real sensor, especially at the higher angles. The ten-pixel model has a total width of 250 µm and a thickness of 250 µm with periodic boundary conditions added to the sides to minimise effects that occur at the edge due to limitations on the simulation size. The reason for using a thickness of 250 µm is to match the real CCPDv3 device and allow for a better comparison to the test beam measurements in chapter 7.4.1. The model is an extension of the three-pixel model previously described in chapter 5.3.1.

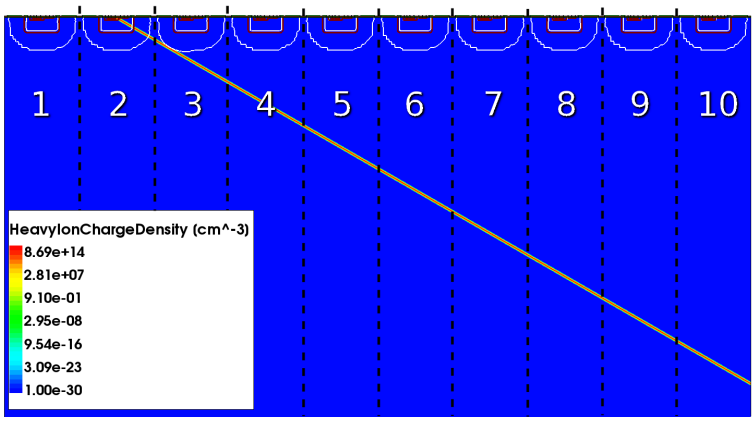


Figure 5.22: An example of a the charge generation due to a MIP traversing the ten-pixel model at an angle of 60°.

Again, the method for generating the MIP signals of the ten-pixel model is the same as the one described in chapter 5.2.6. An example of the charge generation due to a MIP traversing the ten-pixel model is shown in figure 5.22, with the pixels labelled 1-10 from left to right and the MIP entering at the centre of pixel 2. An example of the current pulse shapes for a MIP passing through the 10-pixel model at an angle of 60° is shown in figure 5.23. The hit pixel, pixel 2, has a quick charge collection due to the majority of the MIP's path passing through the depletion region. The next neighbouring pixel, pixel 3, benefits from the MIP path passing through the whole pixel width and being close to the depletion region. Therefore, it collects the most charge. The method of charge collection for the subsequent neighbours, pixel 4 and 5, is diffusion so their collection is slow.

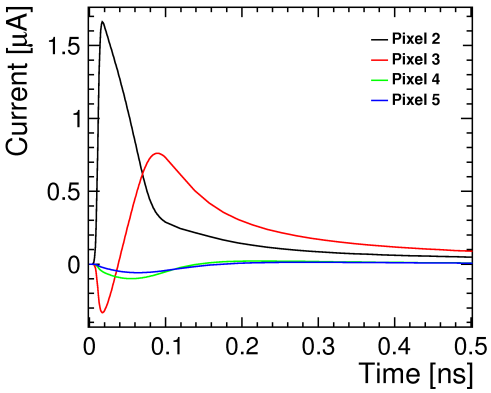


Figure 5.23: Current as a function of time for a track passing through the 10-pixel model at an angle of 60°, for four pixels. Pixel 2 is the hit pixel and pixels 3 to 5 are the subsequent neighbours, as shown in figure 5.22.

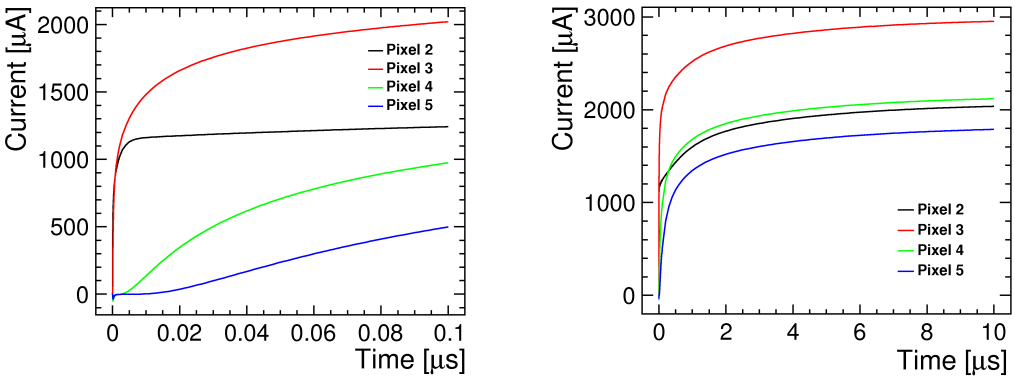


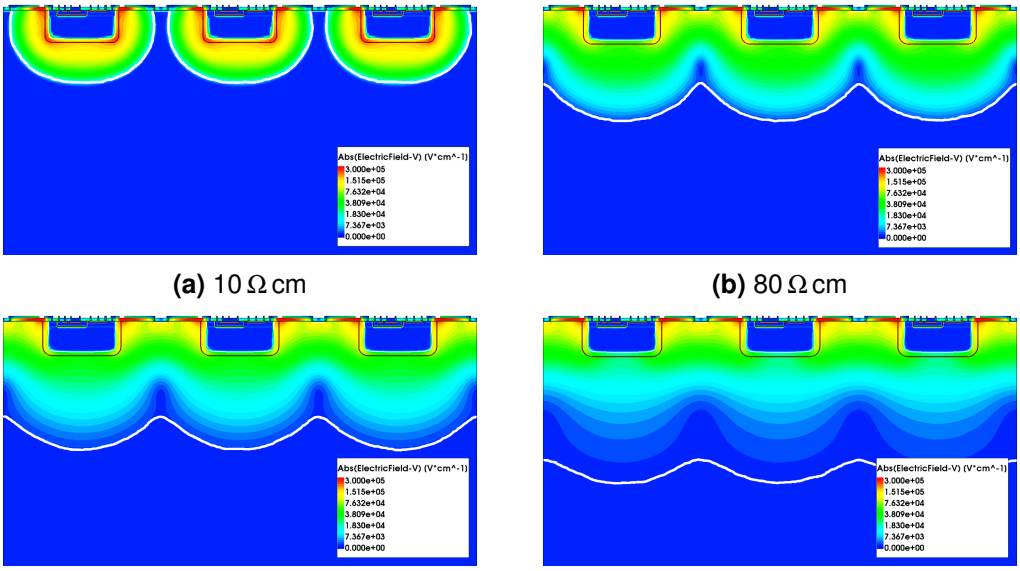
Figure 5.24: The charge collected after (a) 100 ns and (b) 10 µs for a MIP passing through the 10-pixel model at an angle of 60°. Pixel 2 is the hit pixel and pixels 3 to 5 are the subsequent neighbours, as shown in figure 5.22.

5.5 Prospects for improved sensor performance

One method for improving the performance of the sensor is to use a higher resistivity for the substrate. Another method is to apply the bias from the backside by the implantation of a p+ layer. From theory, the larger the substrate resistivity is, the larger the depletion region. Hence, drift will contribute more to the charge collection, resulting in faster charge collection, which is important for CLIC. Another benefit is that the breakdown voltage will increase. These factors are also important for sensors that will operate at the HL-LHC. To corroborate these expectations, a series of simulations were carried out for several bulk resistivity values: 10 Ω cm (current sensor), 80 Ω cm, 200 Ω cm and 1 kΩ cm. Simulations were also performed for the two biasing schemes: topside and backside. This section will highlight a few of the increased performance capabilities that these two methods produce and compare the results. The model used for these studies is the three-pixel model with a thickness of 250 µm.

5.5.1 Electric field

The electric field of a sensor is an important quantity as it can give a better understanding of where the breakdown occurs and the transport direction and magnitude of the carriers. From this, regions within the device that produce fast or slow charge collections can be identified. The value of the electric field for the three-pixel structure was simulated at a bias voltage of −60 V. A comparison of the electric field between the different substrate resistivities is shown in figure 5.25. In all cases, the electric field penetrates the deepest under the deep n-well and has areas of low field under the p+ implants that make up the bias ring. This means that between pixels the charge



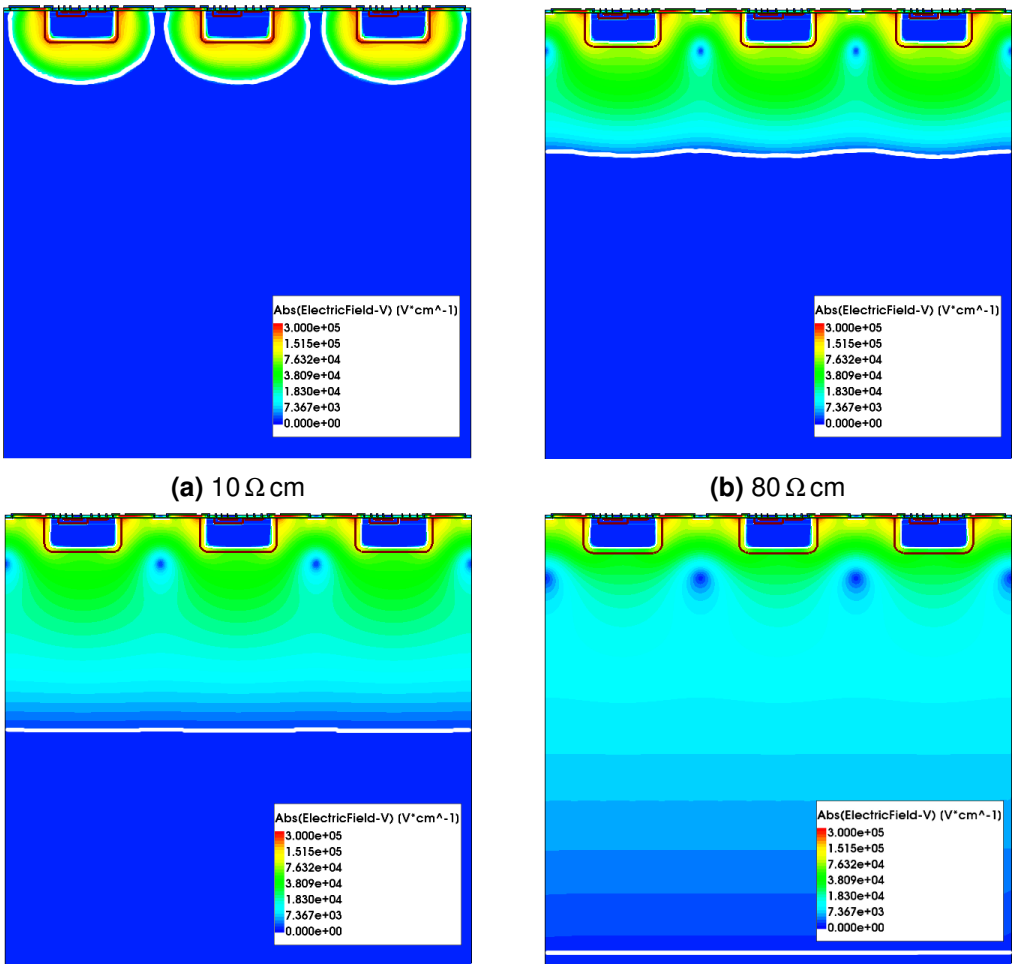
(a) 10 Ω cm **(b)** 80 Ω cm

(c) 200 Ω cm **(d)** 1 kΩ cm

Figure 5.25: Absolute value of the electric field at −60 V for four different values of substrate resistivity: (a) 10 Ω cm, (b) 80 Ω cm, (c) 200 Ω cm and (d) 1 kΩ cm. The white line represents the edge of the depletion region.

collection will be the slowest. This is especially pertinent for the 10 Ω cm case due to very low field values found there. The high field value displayed in yellow is deepest for the 10 Ω cm substrate compared to the higher resistivity values. As the resistivity increases, this area extends laterally, indicating a faster charge collection when the MIP passes in between two pixels compared to lower resistivity models.

In addition to higher resistivities, a further way to improve the performance is to apply bias voltage from the backside instead of the topside. This makes the high values of the electric field extend more towards the bulk which will increase the speed and amount of charge collected. However, at a bulk resistivity of 10 Ω cm there is very little difference. Only at high values of bulk resistivity do the effects become more prominent, especially at 1 kΩ cm, figure 5.26. By taking a cut through the centre of the structure, a comparison of the electric field as a function of depth can be made, figure 5.27. This shows that the higher the resistivity, the greater the difference between the topside and backside bias.



(a) 10 Ω cm **(b)** 80 Ω cm

(c) 200 Ω cm **(d)** 1000 Ω cm

Figure 5.26: Field maps showing the absolute value of the electric field, all to the same scale, at –60 V for biasing from the backside at different bulk resistivity values: (a) 10 Ω cm, (b) 80 Ω cm, (c) 200 Ω cm and (d) 1 kΩ cm.

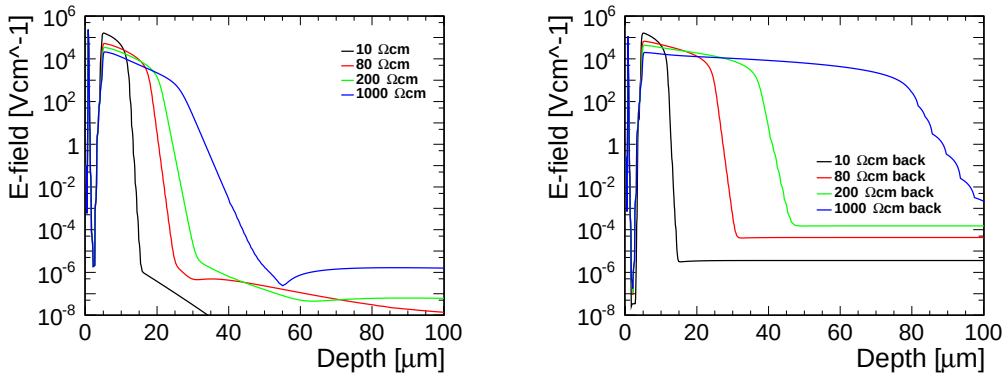


Figure 5.27: A comparison between the absolute value of the electric field as a function of depth for various bulk resistivities and for (a) the top and (b) back biasing schemes.

5.5.2 Depletion depth

As expected, increasing the applied bias and the bulk resistivity leads to a larger depletion region, figure 5.28. To perform a cross-check, the expected values for the depletion depth of a standard diode, shown in dashed lines, are compared. These values are computed from equation 2.9 and assuming an abrupt junction. An abrupt junction is when the doping concentration of a pn-junction is a step function.

There is a good agreement between the simulation and prediction (from equation 2.9) at 10 Ω cm for both topside and backside, with the prediction being slightly larger. However, for the topside bias scheme, as the resistivity increases, so does the difference between simulation and prediction. This is due to the assumptions made in equation 2.9. One such assumption is that it is an infinitely large device consisting of a p-type side and an n-type side, whereas in the simulations the n-type is segmented. The major effect is that the bias is applied from the topside in the simulations and not the backside (as is the case in equation 2.9), which has a large influence on the depth. This results in the depletion depth being greater for backside biasing. Similar to the electric field, the backside bias improves the depletion depth to a greater degree at the higher resistivities, an increase of 55 µm for the 1 kΩ cm model. In contrast to the topside, the backside depletion depth values are much closer to the theoretical values of an abrupt junction.

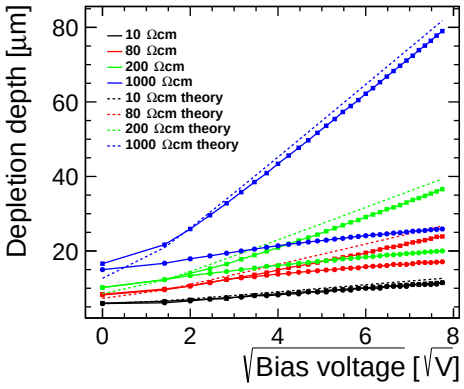


Figure 5.28: The depletion depth as a function of the square root of bias comparing various resistivities and the two bias schemes. The topside bias is represented by the circles and the backside with squares, while the dashed lines are the theoretical values for an abrupt junction.

5.5.3 Leakage current

The leakage current of the three-pixel model at different bulk resistivities is shown in figure 5.29. As the resistivity increases, so does the breakdown voltage. However, the three cases with the larger values for the bulk resistivity show a similar breakdown at around −100 V for topside biasing. This indicates that there is a limiting factor stopping the breakdown voltage becoming larger than −100 V. One possible explanation is that the large electric fields, found at the silicon oxide boundary on the surface of the device, are related to the space between the p+ implant and the deep n-well. By going to higher resistivities this can only be partially compensated for. The greatest difference in breakdown voltage is seen between the 10 Ω cm and the other resistivity models, with a difference of around 10 V. This is important because the operating bias can be increased, producing a larger electric field and depletion region. Hence, a faster and larger charge collection for the sensor, details of which are discussed in chapter 5.5.5.

For the backside biasing scheme, the 10 Ω cm model breaks down at −125 V, while for the higher resistivities there is a large improvement with the breakdown occurring above −800 V. One drawback of the higher resistivities is that there is a larger leakage current. Therefore, there will be more noise. The difference in the leakage current is less than an order of magnitude between 10 Ω cm and 1 kΩ cm and as the voltage increases the difference reduces. This is because the leakage current is proportional to $\sqrt{V}$, as seen in equation 2.21. There is a slight increase in the leakage current for the backside compared to the topside, as shown in figure 5.29, due to the increase in depletion width.

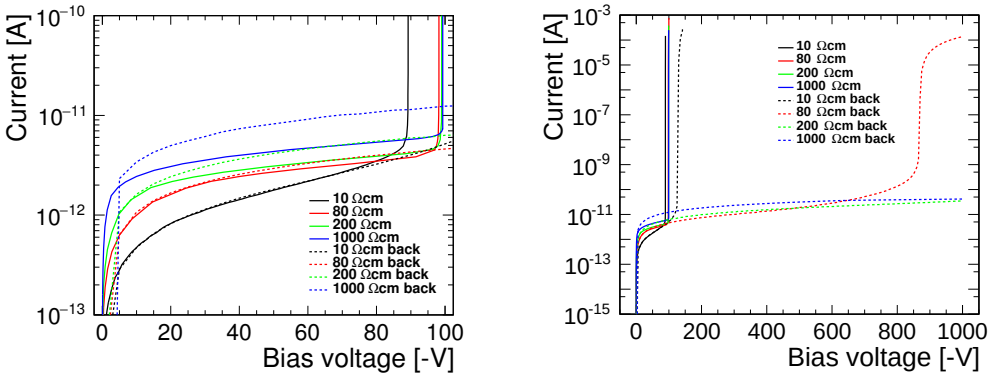


Figure 5.29: Current as a function of bias at different bulk resistivities and for different biasing schemes showing two voltage ranges.

5.5.4 Capacitance

The value of the deep n-well-to-bulk capacitance of the three-pixel model for the four different bulk resistivities was acquired using the same simulation method as described in chapter 5.2.5. This value, shown in figure 5.30, shows smaller capacitances for higher resistivity substrates, resulting in less detector noise. This is expected because, as presented in chapter 5.5.2, the larger resistivities have larger depletion regions which the capacitance is inversely proportional to, equation 2.10. As with the leakage current, there is little difference between the three higher resistivity cases and the largest difference is from 10 Ω cm to the other resistivity models. The kink of the 80 Ω cm and the 200 Ω cm curves is produced when the depletion region reaches the edge of the structure and is an artefact of the simulation, as detailed in appendix B. In both biasing schemes, the improvements seen at higher resistivities are magnified when biasing from the backside. Since the breakdown voltage will be higher for higher resistivities and backside biasing, the operating voltage will be higher. As a result, the operating capacitance will be lower than the values shown here.

5.5.5 Response to a Minimum Ionising Particle

The method for generating the MIP signals in the three-pixel model is the same as the one described in chapter 5.2.6. The current pulse shape from the MIP signal is shown in figure 5.31. The peak height decreases for larger resistivities and the peak times are approximately the same (10–20 ps). At 0.5 ns the larger resistivities have larger current values. The collected charge of the simulations is found by integrating the current over time; the resulting values are shown in figure 5.32. As indicated by the current pulse shapes, the lower resistivities initially have a quicker charge collection in the sub-nano

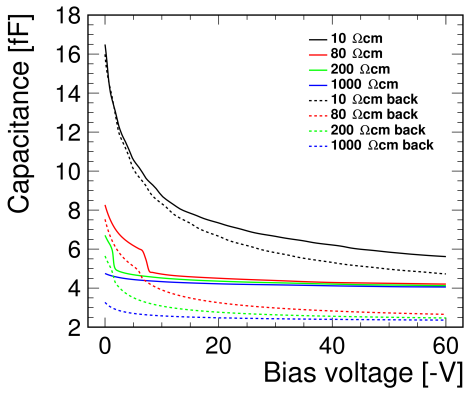


Figure 5.30: Deep n-well to bulk capacitance as a function of voltage comparing various resistivities and the two biasing schemes.

second range. This is a consequence of the lower resistivity models having a larger electric field value in the depletion region, but after 0.5 ns the lower resistivity models' smaller depletion region means that the higher resistivities begin to collect more. The 1 kΩ cm model has the highest collected charge, with around 700 e^- more charge collected than the 200 Ω cm model after 100 ns and 10 µs. Initially the 10 Ω cm model collects the least charge from 0–25 ns, due to having the smallest depletion region, but then collects more than the 80 Ω cm model after 100 ns. This is because diffusion is now the dominant method of charge collection and the 10 Ω cm model collects more charge due to diffusion, this will be discussed and illustrated in chapter 5.6. After 3.5 µs the 80 Ω cm model again collects more charge than the 10 Ω cm model, which is due to more recombination and charge sharing to neighbouring pixels in the 10 Ω cm model.

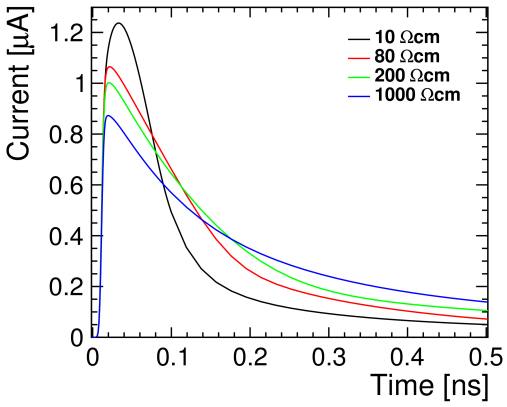


Figure 5.31: The current pulse shape of the simulated MIP passing through the centre of the central pixel cell for different bulk resistivity values.

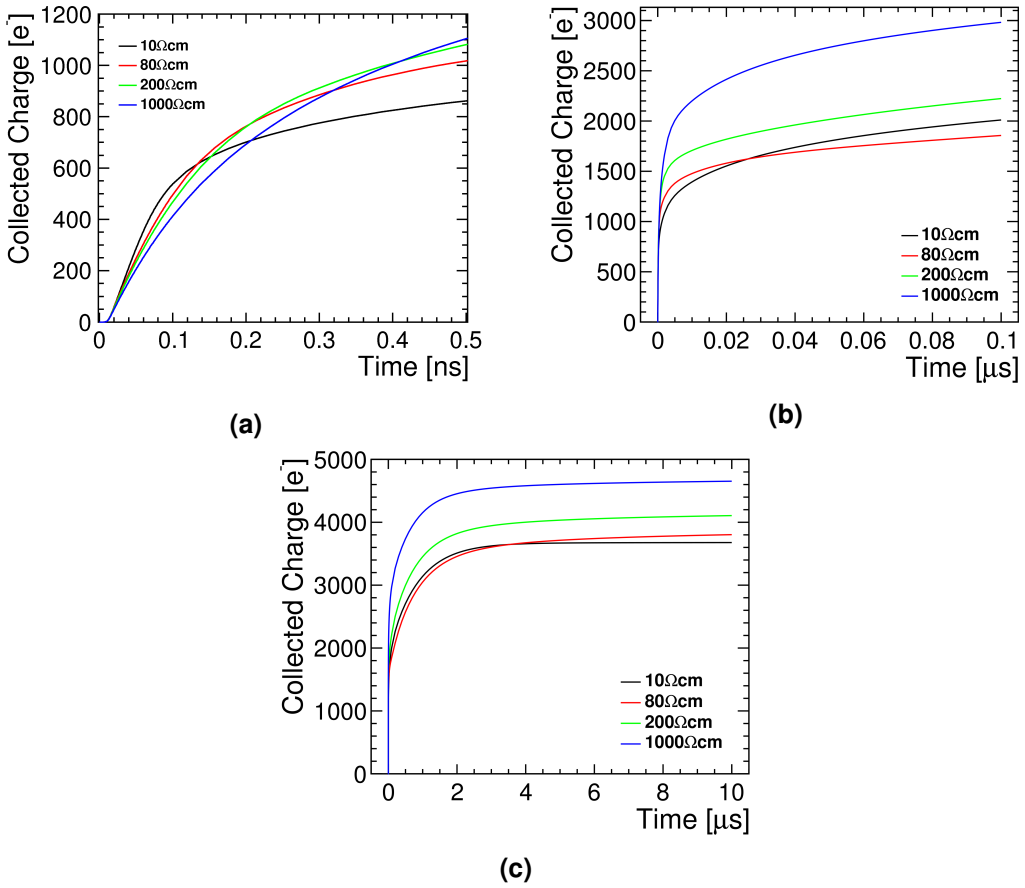


Figure 5.32: The total charge collected for the three-pixel model at different bulk resistivity values over different time scales: (a) 0.5 ns (b) 100 ns and (c) 10 µs.

To understand the timing response of the HV-CMOS sensors, it is important to estimate the fraction of charge collected by drift as it produces a fast signal. From the depletion depths shown in figure 5.18(b), an estimate of the amount of charge collected by drift can be made. In the 10 Ω cm model the depletion depth is 11.4 µm for a bias voltage of −60 V, resulting in a drift charge of around 900 e^-, around one quarter of the total charge collected after 10 µs. In the 1 kΩ cm model the depletion depth is 25.9 µm, resulting in a drift charge of around 2000 e^-, which amounts to under one half of the total charge collected after 10 µs. So, larger a resistivity means a larger proportion of the charge is due to drift hence, a faster signal.

As expected from the significant increase in the depletion region throughout the sensor depth between the topside and the backside biasing schemes, the current pulse shapes are much wider (figure 5.33). Consequently, the charge collection is also greater for the backside after 0.5 ns and remains like this until 10 µs (figure 5.34). As shown in the other sensor properties, the largest improvement between the two is seen in the 1 kΩ cm model with a difference of 1400 e^- after 0.5 ns, a difference that increases with

time. From figures 5.25 and 5.26, the electric fields for the 10 Ω cm model are almost identical, leading to a very similar charge collection between the two schemes.

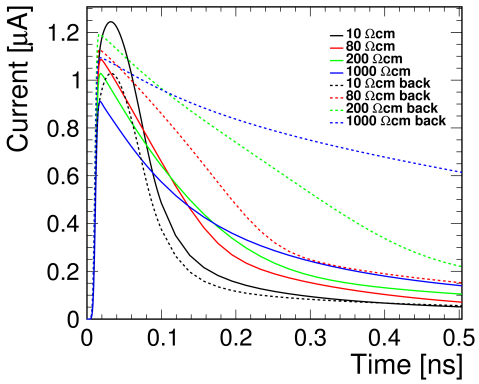


Figure 5.33: A comparison of the current pulse shapes for various bulk resistivities and the top and back biasing schemes for a MIP passing through the centre of a pixel cell.

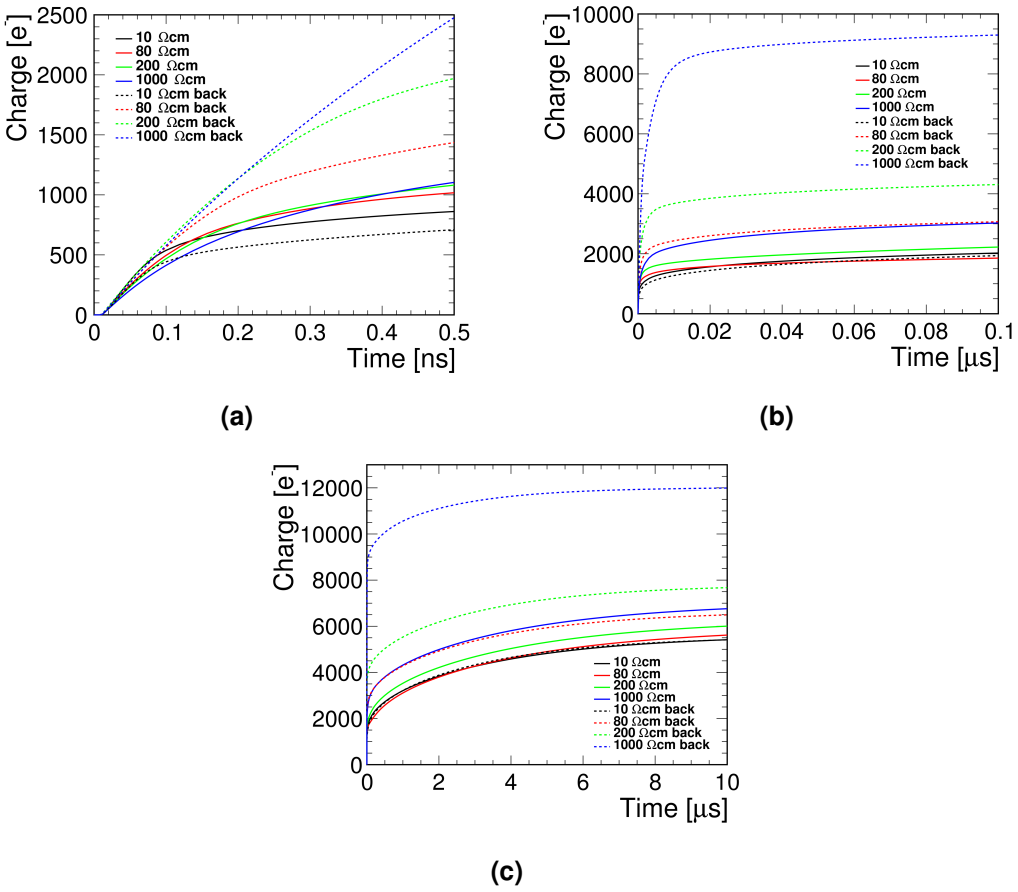


Figure 5.34: The charge collected for topside and backside biasing schemes of various bulk resistivities after (a) 0.5 ns, (b) 100 ns and (c) 10 μs.

5.5.6 Angled simulations

For all the previous plots the tracks were of perpendicular incidence. One concern of using higher resistivities is less of the charge will be collected from diffusion, resulting in a smaller cluster size. Therefore, the position finding algorithm will not be as effective[1]. By looking at the response to a MIP through the centre of a pixel at angles of up to 80°, a better understanding of the cluster size is possible. This is done using the 10-pixel model described in chapter 5.5.6 and has an arbitrary detection threshold of 500 e^- for a pixel to be part of the cluster. The collection times would vary depending on how much of the particle path went through the depletion zone of each pixel. If the particle passed through most of the depletion zone then that pixel would reach the 500 e^- threshold in $\approx$ 10 ns. If this was not the case and the majority of charge was due to diffusion, then the collection time would be $\approx$ 100 ns.

The total charge collected from all pixels above the threshold for various angles is shown in figure 5.35. At all angles the 1 kΩ cm model collects the most charge for topside biasing. Below 60°, the three lowest resistivities have a similar value to each other. Above this angle, the 10 Ω cm model matches the 1 kΩ cm vales. In the topside biasing case the higher the resistivity the more total charge is collected. The drop at 80° seen for the 1 kΩ cm model is not physical and is due to the model only having 10 pixels and the charge flowing to the pixel on the other side via the PBC.

The reconstructed cluster x-width (the number of pixels within the cluster in the x-direction) is shown in figure 5.36. At 0° there is no difference in the cluster x-width for either a higher resistivity or biasing from the backside, all models show a value of one. The general trend for the topside biasing is that all the resistivities are very similar and only the 1 kΩ cm model shows a slight improvement from the 10 Ω cm model.

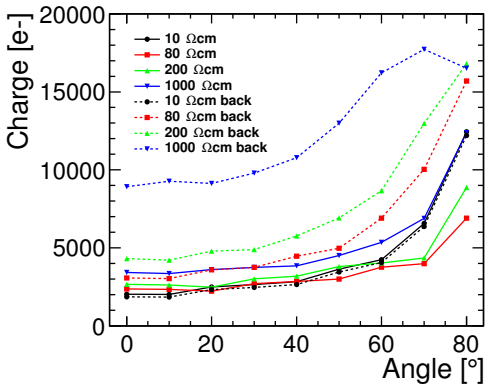


Figure 5.35: The total charge collected from all pixels above the threshold as a function of angle.

[1] This is not always the case and depends on the cluster finding algorithm.

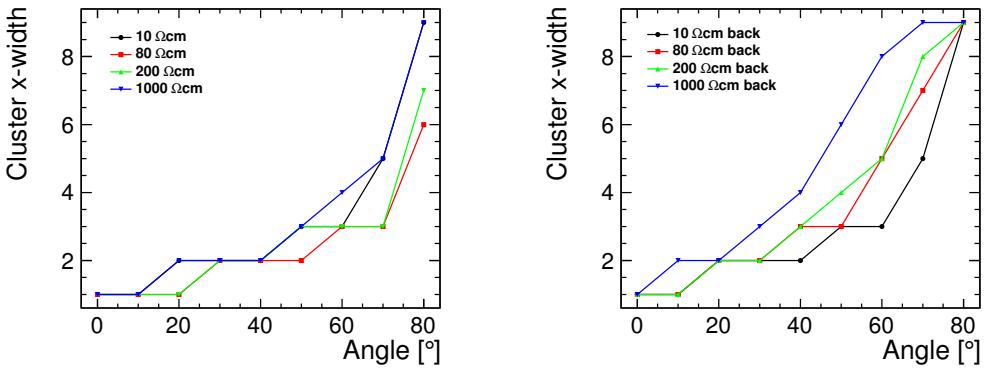


Figure 5.36: The reconstructed cluster x-width as a function of angle comparing various bulk resistivities and the: (a) topside and (b) backside biasing schemes.

the higher resistivities, especially 1 kΩ cm, having a larger cluster x-width even at lower angles. The trend of increasing angle resulting in increasing cluster x-width is not seen from 70° to 80°. This is due to the model only having 10 pixels and the MIP entry point being in pixel 2 hence, the largest cluster x-width is 9.

The values obtained here are very sensitive to the threshold values used. They are only meant to give an indication of the effects of using a higher resistivity and biasing from the backside at various angles. More detailed studies using a complete simulation of the readout will be needed along with beam test measurements to determine the impact of the different sensor types.

5.6 Comparison of charge collection methods

To better understand how the signal is collected and the relative contribution of the two collection methods, the path of the MIP was split into two distinct regions: “drift” and “diffusion”, shown in figure 5.37. The “drift” path is defined as only the charge deposited in the depleted region of the sensor, figure 5.37(b), while the “diffusion” path is defined as only the charge deposited in the non-depleted bulk, figure 5.37(c). Transient simulations were run to produce the current and collected charge of the two paths. To check that the splitting of the MIP path was done correctly, the current and charge from the two separated paths were added together, labelled “sum”, and compared to the MIP passing through the whole structure, “all”. A first look at such a scenario will be done for a resistivity of 10 Ω cm. Later in this section, simulations of different resistivity values will be examined. In all cases, the MIP passes through the central pixel and only the charge collected on this pixel is studied.

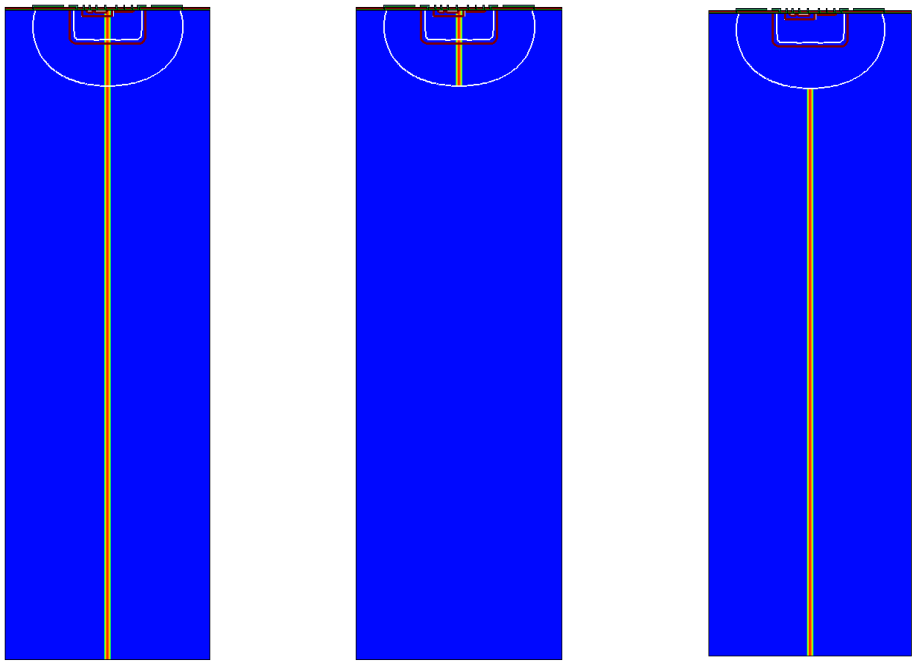


Figure 5.37: The different paths the MIPs take: (a) when the MIP passes through the whole thickness of the structure, “all”; (b) the MIP passes through just the depletion region, “drift”; and (c) the MIP passes only through the non-depleted bulk, “diffusion”. The white line represents the edge of the depletion region.

The current produced by the different paths is shown in figure 5.38. The “drift” current pulse shape has a comparable pulse height and position to the “all” curve while the “diffusion” current pulse shape is much lower in comparison, as expected. The reason that the “drift” current is not identical to the “all” current is that a small amount of charge, deposited just outside of the depletion region, contributes to the “all” peak height. This can be seen by the fact that the “diffusion” curve has a small peak in such a short time frame, 0.1 ns, and that it overlaps with the main peak. However, there is a small discrepancy between the “all” and “sum” curves, with the “all” MIP path producing a larger current. This could be due to the induced charge created by the MIPs affecting the electric field in different ways due to them not having identical paths. Nevertheless, the difference is very small at less than 1%.

The very small difference between the “all” and “sum” curves seen in the current is carried forward to the collected charge but the difference becomes even smaller (less than 1% at both 100 ns and 10 µs - see figure 5.39). The “drift” signal is collected very quickly (collecting 700 e^- in approximately 0.3 ns) but plateaus once around 900 e^- are collected, doing so in around 5 ns. After 50 ns, the “diffusion” signal becomes the dominant method of charge collection and after 10 µs contributes to around 75% of

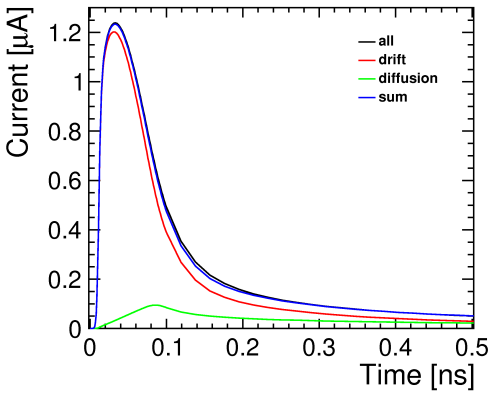


Figure 5.38: The current induced on the collection contacts as a function of time for the various MIP paths. The blue line is the sum of the drift and diffusion current values.

the total charge collected. The value of 75% should be considered as specific to the simulated structure. This is because for the real sensor, with a thickness of 250 µm, the amount of "drift" charge would stay the same but there would be a larger contribution from the "diffusion" part. However, this only plays a role if the sensor amplifier has a long integration time which is not the case for the fast collection necessary at CLIC. A way to increase the amount of fast (drift) signal would be to create a larger depletion region which can be done in three ways: by an increase in the bias voltage, an increase in the bulk resistivity or biasing from the backside. As the maximum bias voltage for this sensor technology is limited due to the breakdown, this only leaves the option to increase bulk resistivity and/or bias from the backside.

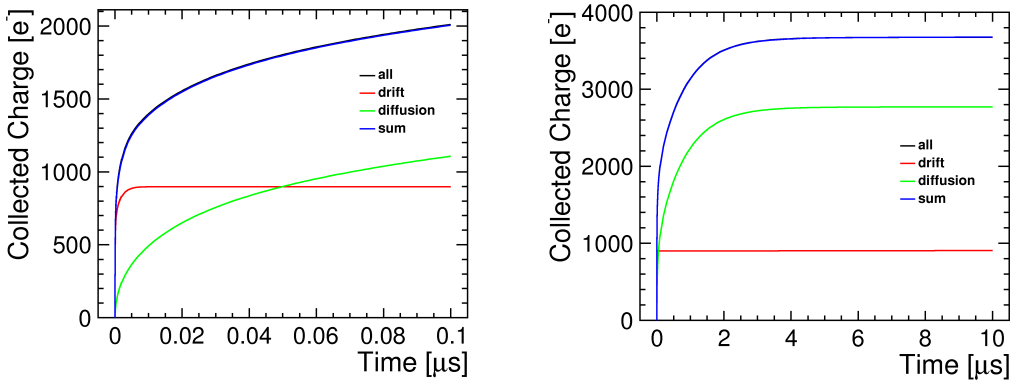


Figure 5.39: The charge collected for the separate MIP path segments over different time scales: (a) 100 ns and (b) 10 µs. The blue line is the sum of the collected charge from drift and diffusion.

Performing the same separation of the MIP path for different resistivities and comparing can lead to a better understanding of the exact nature of the charge collection for each scenario. The depletion depths will play an important role in the amount of drift and diffusion charge collected. The depth values at each resistivity are: 11.5 µm at 10 Ω cm, 17 µm at 80 Ω cm, 20 µm at 200 Ω cm and 25.9 µm at 1000 Ω cm.

The currents seen by the "drift" and "diffusion" MIP paths are shown in figure 5.40. For the "drift" path, the lower resistivities have the higher peaks, which is due to the larger electric field value within the depletion region, as was shown in figure 5.25. For the higher resistivities, on the other hand, they have a broader peak due to the depletion region being larger. This pattern is also reflected in the current pulse in figure 5.31 when the MIP path passes through the whole structure. In the "diffusion" curve, again, the lower resistivities have a higher peak but now all at different times, getting later as resistivity increases. This is because the electric field near the edge of the depletion region is higher for lower resistivities, so charge that diffuses into the depletion region is quickly collected. Moreover, the depletion region is larger for the larger resistivities, thus the charge from the "diffusion" path will take longer to reach the collection electrode.

The charge collection of the "drift" MIP path within 2 ns displays the same trends as those seen in the current pulse shapes: the lowest resistivity initially collecting the most then is quickly overtaken by the higher resistivities. In all cases, they collect the deposited charge and plateau after 10 ns. This is not the case for 1 kΩ cm, which still shows some increase in charge when the integration time becomes long. Since the 80 Ω cm and 200 Ω cm models do plateau, on the 100 ns time scale, but then show a rise again on the 10 µs time scale, it is assumed that this rise is due to the leakage

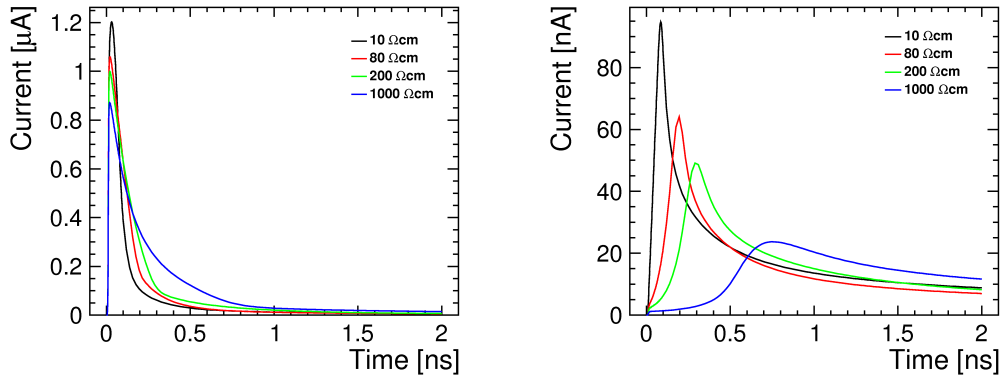


Figure 5.40: The current pulse shape for the two separate MIP paths (a) drift and (b) diffusion for different resistivity values after 2 ns.

current. This can also explain why the 1 kΩ cm model has the steepest line as the higher the resistivity, the larger the leakage current. This is in accordance with equation 2.21, where the leakage current depends on depletion width.

In the case of the “diffusion” MIP path, the 10 Ω cm collects most charge initially and even remains the largest after 10 µs. This is attributed to the fact that the “diffusion” path is the largest for the 10 Ω cm model. However, this does not explain the behaviour of the remaining resistivities. As expected from the current pulse shapes, the lower resistivities collect most charge initially but then as more time passes this is reversed. This may be due to a number of factors: the larger mobility in the higher resistivities, the larger recombination in lower resistivities and/or an artefact of the simulation and the unphysical MIP paths.

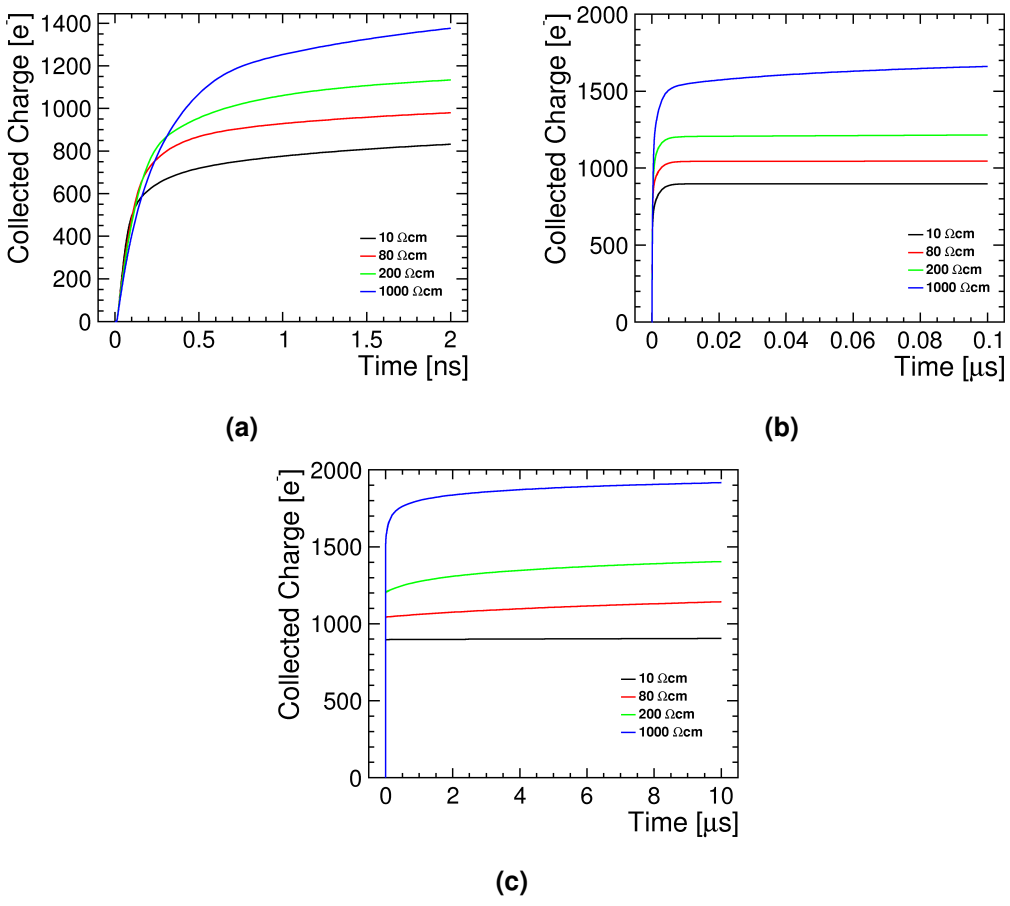


Figure 5.41: The total charge collected for the “drift” MIP path at (a) 2 ns, (b) 100 ns and (c) 10 µs for different resistivity values.

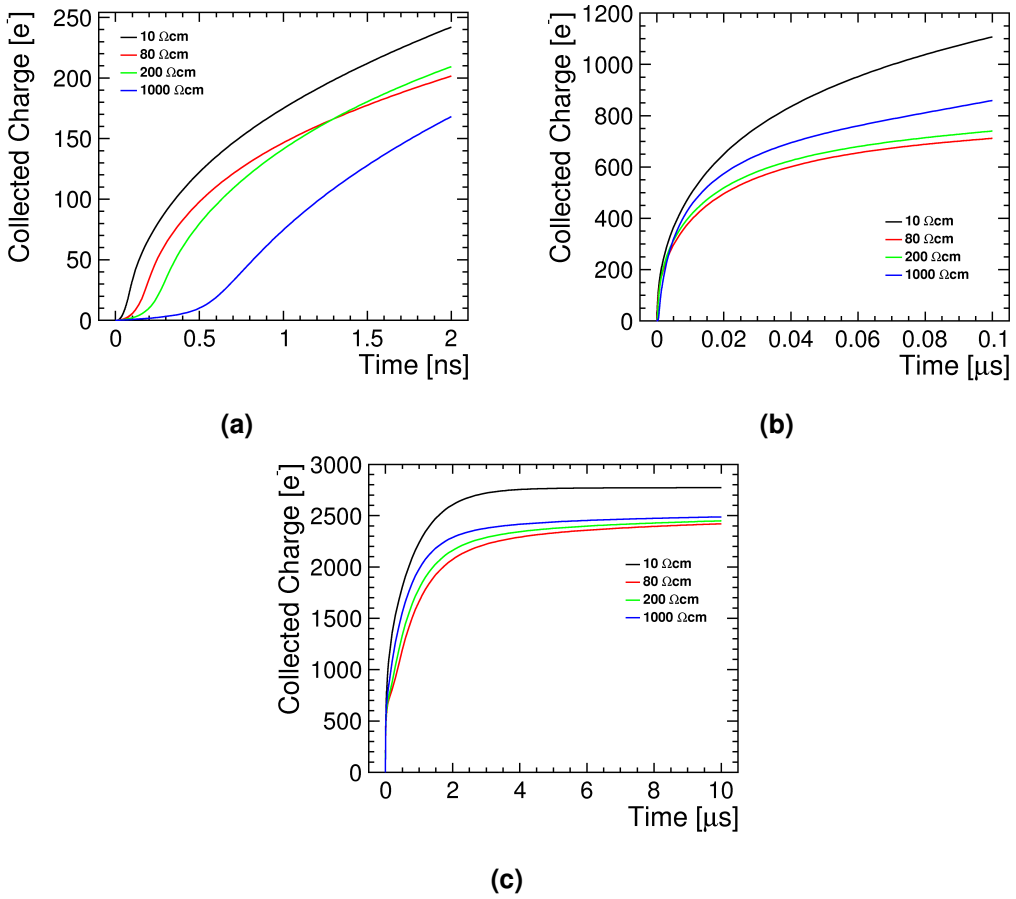


Figure 5.42: The total charge collected for the "diffusion" MIP path at (a) 2 ns, (b) 100 ns and (c) 10 µs for different resistivity values.

5.7 Summary

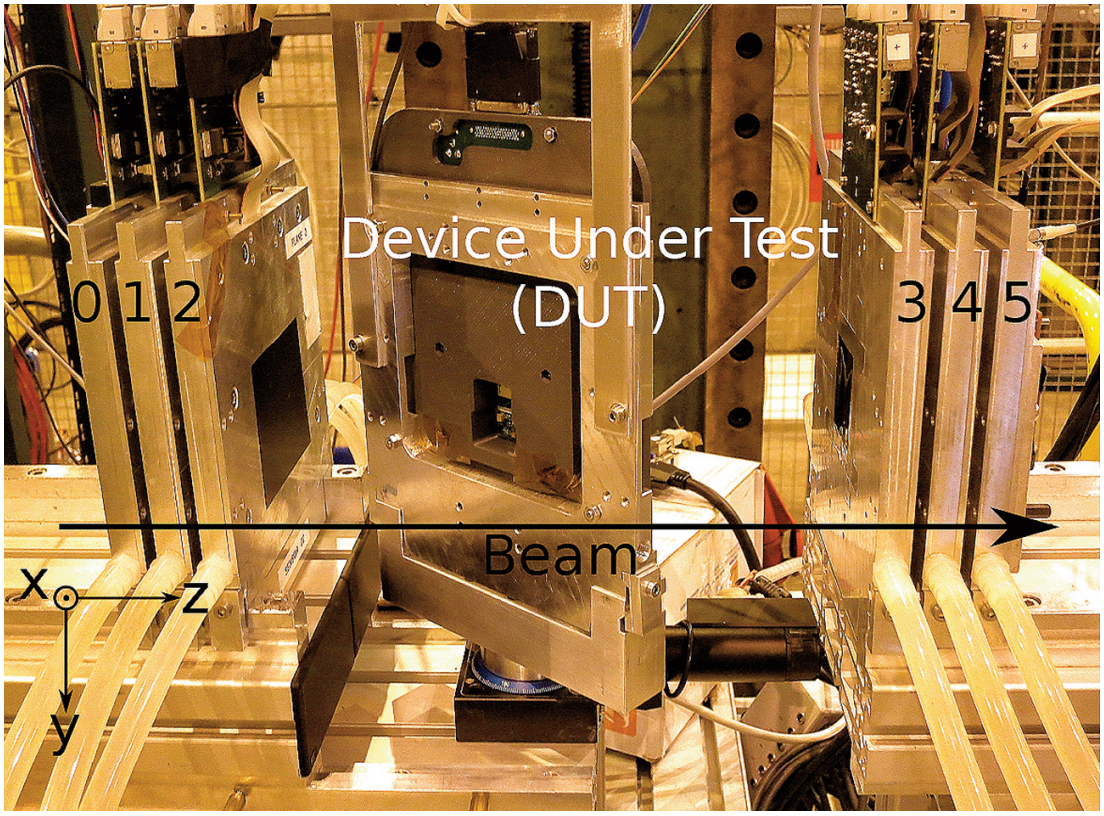
In this study three TCAD simulation models of the CCPDv3 HV-CMOS sensor have been presented. A comparison between the full 2D model, the reduced 2D model and the reduced 3D model has been conducted. Between all three models the difference between the simulated sensor properties is less than 15% in leakage current, breakdown voltage and charge collection after 100 ns (the simulated time at which the charge collected contributes to the signal). However, there is a difference of 40% in the capacitance results. This indicates that the 3D simulations are better suited to extracting the capacitance as the results are closer to the data values and the simulation time is within an acceptable limit, $\approx$ 1 day. For the charge collection simulations, the full 2D results are reasonably close the reduced 3D results. The difference of 15% between the 2D and 3D is acceptable due to the fact that 3D simulations of a MIP use drastically more memory ($>$ 64 GB) and time ($>$ 1 week), compared to a memory usage of $<$ 4 GB and a time of $<$ 3 hours.

In this study there was no full 3D model simulated due to resource constraints, so the effects of using the full 3D model cannot be known. Nevertheless, the difference between the reduced 2D and full 2D gives an insight into the results for the full 3D model. For the charge collection, the full 2D model prediction is found to be larger than the reduced 2D model. Assuming a similar pattern when comparing the full 3D model with the reduced 3D model, this could lead to a smaller difference between the full 2D model and the full 3D model than those found by comparing full 2D and reduced 3D predictions.

In the simulations of the full 2D three-pixel model, there is a good agreement with simulations and the measurements of the leakage current, breakdown voltage and the depletion depth. For the charge collection, once 100 ns have passed, the charge collection across the whole device is approximately uniform for every MIP entry position.

By comparing the results of higher resistivity and biasing from the backside, the possible prospects for improved sensor performance were found. In all the results shown, going to a higher resistivity improves the performance. There was also an improvement in all the results when biasing from the backside, especially for the 1 kΩ cm model. This means that if the backside bias was to be implemented, a higher resistivity would maximise the improvement. However, the improvements were not seen for the leakage current, where a larger leakage current of up to 2 times was observed due to the larger depleted volume. The improvement in sensor performance is also seen in the charge collection at different incidence angles where higher resistivity and biasing from the back improves the total charge collection and the cluster x-width. This is of importance for the position finding algorithms. However, more detailed measurements would need to be done to confirm this. One drawback of backside biasing is that an extra processing step has to be applied on the backside which adds additional complexity and cost. Furthermore, biasing from the backside increases the complexity of the wiring for the voltage lines, potentially increasing the material budget.

Splitting the MIP path into the two types of charge contribution showed that, as expected, collection by drift is very fast but is limited by the depletion region size. While diffusion is slower, it becomes the dominant method of charge collection after 50 ns for the 10 Ω cm model. Comparing the different resistivities, one sees that the higher values initially collect charge slower for both the “drift” and “diffusion” paths. This is very quickly reversed for “drift”.



The AIDA beam telescope in the H6 beam line at the CERN SPS North Area, showing the six telescope planes and the device under test mounted onto arotation stage. It also defines the coordinates of the system.
(see page 127)

CHAPTER 6

Calibration and characterisation of sensor assemblies

Prior to measurements conducted in the beam test, characterisations and calibrations are conducted to determine optimal operating settings for the sensor and identify any potential problems with the assemblies, e.g. high leakage current or non-responsive regions. These calibrations are used to convert the raw measurements into physical units, typically the number of generated electron-hole pairs. This is done by determining the relationship between the threshold digital-to-analogue converter (DAC) setting of the discriminator, the time-over-threshold (ToT) measurement, defined in chapter 2.3, and the energy deposited in the sensor. This chapter describes the various laboratory measurements conducted to achieve the aforementioned goals. Two types of assemblies will be investigated: planar sensors bump-bonded to the CLICpix readout chip and HV-CMOS sensors capacitively coupled to the readout chip.

6.1 Set-up

Measurements were conducted in the lab at CERN, with the various boards placed inside lead shielding. The readout system is custom built and uses a SPARTAN6 FPGA board with a modular interface card, the µASIC [79]. The chips are mounted on a carrier board from which wire bonds connect to the assembly. A picture of the three boards connected together is shown in figure 6.1. A close up of both types of assembly is given in figure 6.2, showing the drops of glue on top of the planar sensor. The high-voltage to the chip was supplied by a Keithley 2450 source measure unit (via a LEMO cable to the interface board). The FPGA board was connected and controlled by ethernet cable to a computer from where the data acquisition (DAQ) was controlled and scripts could be used to carry out the measurements.

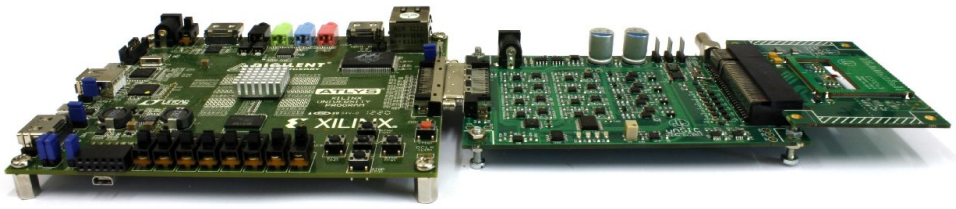
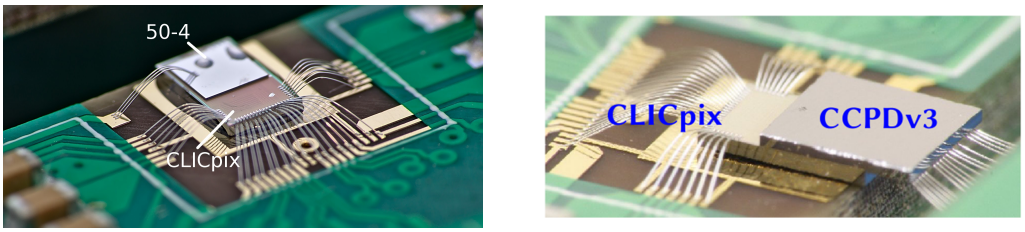


Figure 6.1: A picture showing the three boards used in the readout system. On the left is the FPGA board, in the middle is the µASIC interface board and on the right is the chip carrier board holding the assembly [79].



(a) (b)

Figure 6.2: A close up of a (a) planar sensor assembly and (b) a capacitively coupled assembly, showing the wire bonds on both sides and the carrier boards.

To correct for the threshold dispersion on the readout chip caused by uncertainties in the fabrication process, an equalisation was used. This uses a 4-bit DAC in the discriminator, called the calibration DAC, to adjust the threshold for each pixel separately. This is different from the 12-bit DAC in the periphery of the readout chip that controls the global threshold. By design for the 12-bit DAC, one DAC step corresponds to approximately 10 e^- [52]. To avoid confusion, the calibration DAC uses hexadecimal. The equalisation finds the global baseline DAC value for the whole matrix, the baseline was defined in chapter 2.3. Then it finds a baseline value for the lowest (0) and highest (F) value of the 4-bit calibration DAC for each pixel. From this, the equalisation script interpolates the characteristics of the baseline in response to a change in calibration DAC value for each pixel. These characteristics are then used to set a calibration DAC value for each pixel that equalises the effective threshold. The equalisation process is described in more detail in [50, 80]. A graph showing the threshold dispersion for the equalised chip as well as the lowest and highest calibration DAC values is shown in figure 6.3. The threshold dispersion before calibration is 19.1 DAC for calibration DAC=0, 15.8 DAC for calibration DAC=F and 3.2 DAC for the equalised matrix. These values correspond to the sigma of a Gaussian fit and is called the noise RMS.

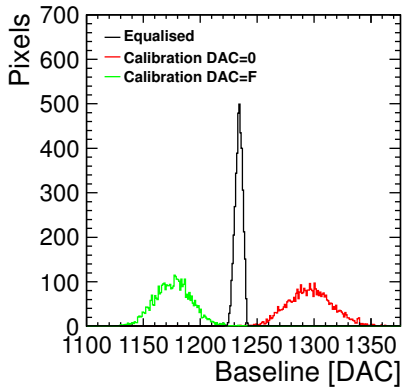


Figure 6.3: A graph showing the dispersion of the 12-bit global threshold before equalisation, for two calibration DAC values of 0 and F, and after equalisation.

The global baseline, the baseline over all pixels, is obtained from the mean of the Gaussian fit to the threshold dispersion after equalisation. In order to be above the noise level, the operating threshold had a value of 120 DAC subtracted from the baseline (a decrease in DAC values means an increase in threshold, due to the polarity of the chip, negative). A clock frequency of 20 MHz and a shutter time of 10 ms were set on the CLICpix. A constant discharge current for the ToT measurements was set to an I_{krum} value of 25. The I_{krum} current is used to control the discharge rate of the feedback capacitor and influences the return to baseline for the discriminator and hence the ToT measurement, more details can be found in [81]. No pixels were masked initially so that a handle on the quality of the pixels could be acquired. The set-up for both the planar and capacitively coupled assemblies are identical except for the threshold and voltage used to power the chips.

6.2 CLICpix planar sensor assemblies

One option for the CLIC vertex detector is a hybrid assembly that connects a planar sensor to the readout ASIC, details of such assemblies were given in chapter 4.3.2. To determine the suitability of these assemblies, the quality of the fabrication and the quality of the bump-bonding process were assessed with laboratory measurements. For this study, four usable assemblies were produced: three at a sensor thickness of 200 µm, labelled 14, 31 and 43, and another assembly for which the sensor was thinned down to 50 µm, labelled 50-4.

The operation thresholds used and the noise RMS values for each assembly are shown in table 6.1. Looking at the noise RMS values of each pixel, figure 6.4, it can be seen that a few pixels have higher noise RMS values, giving a possible indication to problematic areas in the assemblies. All the measurement conditions are identical unless stated otherwise. The threshold, bias voltage and clock frequency setting used for calibration were set to match those that were used in the beam tests. In the source calibration measurements, the threshold for all assemblies was reduced by a DAC value of 60. Therefore, to match this, so were the thresholds of the test pulse calibrations. The threshold at which the measurements were performed was at a value of 60 DAC less than the ones shown in table 6.1. The reason for this was that at the values given by the equalisation, described in chapter 6.1, it was difficult to distinguish the signal from the noise. The threshold was then incremented to higher values (lower DAC values) until a signal could be clearly seen. The test pulse calibrations were done at the same threshold. This ensures that the source ToT measurements can be turned into physical units using the test pulse curves.

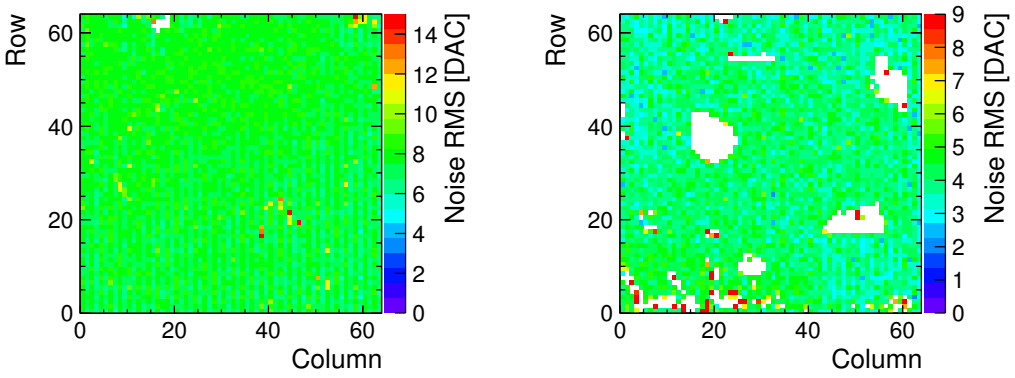


Figure 6.4: The noise RMS in DAC after equalisation for assemblies: (a) 31 and (b) 50-4. The pixels with a noise RMS value of 0 are dead pixels.

6.2.1 Current-voltage characteristics

The first set of tests to be carried out on the assemblies is to check that all the wire-bonds are functioning and there are no shorts on the carrier board. This is done by setting the voltage regulators, that are used to supply the various electronics of the two chips, and check that the correct voltages are set and the currents are not too high. After this, a check for the voltage that is used to bias the sensor is done by monitoring the current as a function of applied bias in a dark environment. The resulting curves are shown in figure 6.5, and from them the breakdown voltage can be acquired, shown in table 6.2. The breakdown voltage is deemed to be the voltage at which a sharp rise in current can be seen. For assemblies 14 and 43 this occurred earlier than assembly 31. The operating voltage was chosen to be at a level where the leakage current was sufficiently low, below a few tens of nA, and above the estimated depletion voltage, $\approx$–35 V. This depth was estimated from capacitance-voltage measurements on test diodes, which were later supported by beam test measurements.

Table 6.1: The assemblies investigated in this study and their operating thresholds.

Assembly	Sensor thickness [µm]	Operation threshold [DAC]	Noise RMS [DAC]
14	200	942	6.97
31	200	1111	7.61
43	200	917	6.64
50-4	50	962	3.98

In the case of the 50-4 assembly, the leakage current was very high and higher than expected, so only a small voltage could be applied. An estimate of the depletion voltage, assuming it had a similar resistivity as the thicker assemblies, was given as ≈–1 V. Initially, assembly 14 could be operated at –50 V, as seen in the figure 6.5, but unfortunately could no longer be biased to high voltages due to very large leakage currents, so its results in the following section will be shown for 0 V and calibrations were not performed. This happened after the sensor was held at –50 V for a long time during the beam tests.

6.2.2 Bump bond quality

The quality of the bump bonding needed to be determined to gauge how successful the new technique is, bonding with such a fine pitch on a chip-to-chip level has not been attempted prior to this. For this study, two types of problem pixels were identified by looking at a ^{90}Sr-hitmap shown in figure 6.6. These maps were taken over 1×10^6 frames for the 200 µm thick assemblies and 100k frames for the 50 µm thick assembly. One possible problem that may arise is that the connection between the two chips is not made or a bump is missing. This can be easily identified by looking at the response of the matrix to a radioactive source. If a connection is not made, then the charge

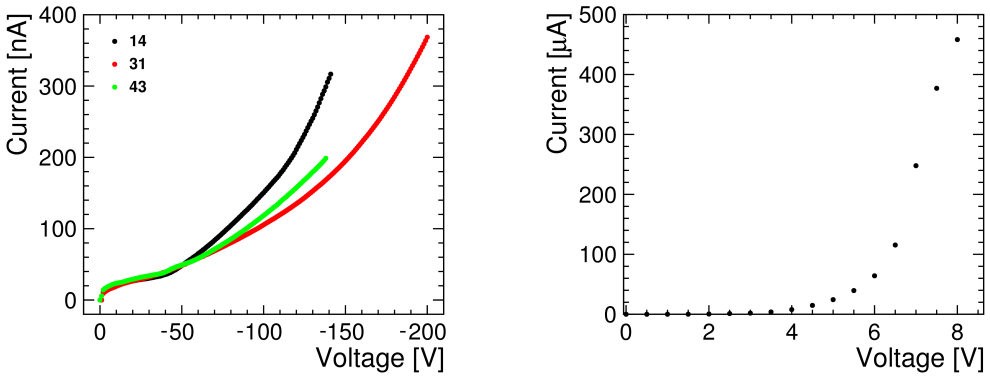


Figure 6.5: The current as a function of applied bias for the (a) three 200 µm assemblies and (b) the 50-4 assembly.

Table 6.2: The operating bias and breakdown voltages of the assemblies.

Assembly	Bias voltage [V]	Breakdown voltage [V]	Depletion voltage [V]
14	-50 V	-140 V	-35 V
31	-50 V	-200 V	-35 V
43	-50 V	-140 V	-35 V
50-4	-2 V	-6 V	-1 V

deposited by an ionising particle cannot make it from the sensor to the readout chip, producing a pixel that does not respond. Another easily identifiable defect is pixels that have a very low response arising from either a weak bond connection or having only a capacitive connection between the two chips. To calculate this a 1D histogram of the number of hits was plotted and fit with a Gaussian to obtain the mean and standard deviation, figure 6.7. Pixels which were more than 3σ lower than the mean response of the pixel matrix, but still show non-zero response, were judged to be weakly responding. Three of the four assemblies show sizeable regions which are not as responsive as the rest of the matrix, while for assembly 31 this is only a small region.

The defects seen in the hitmaps are related to the bump-bonding process, as images of the sensor and readout chip before the completion of the flip-chip process also show similar patterns. An example is shown in figure 6.8, where the lighter areas seen on the sensor and readout chip are seen in the hitmap as weakly responding and pixels which are dark on the readout chip are non-responsive. These are a mixture of problems

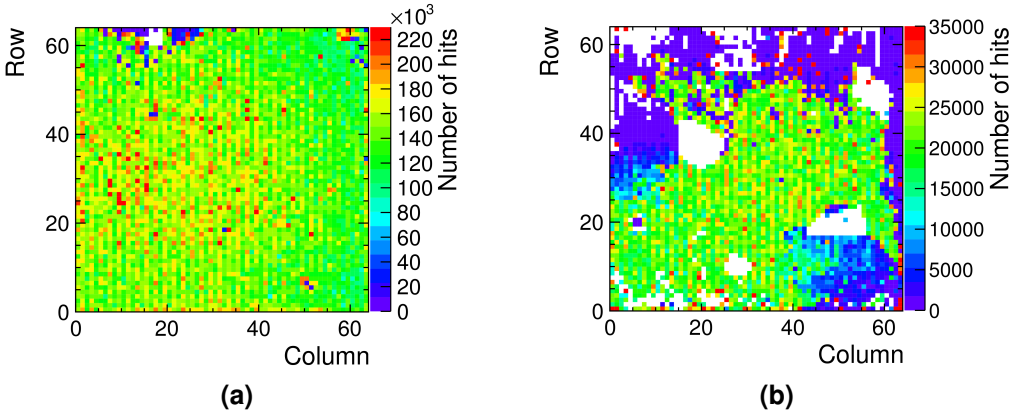


Figure 6.6: The ^{90}Sr-hitmaps of two planar assemblies: (a) 31 and (b) 50-4.

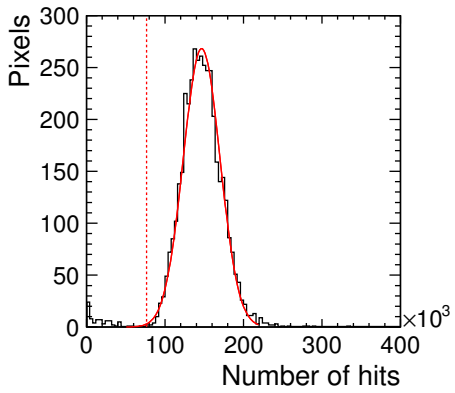


Figure 6.7: The ^{90}Sr hit response of planar assembly 31 showing the Gaussian fit and the cut to define the weakly responding pixels, dashed vertical red line.

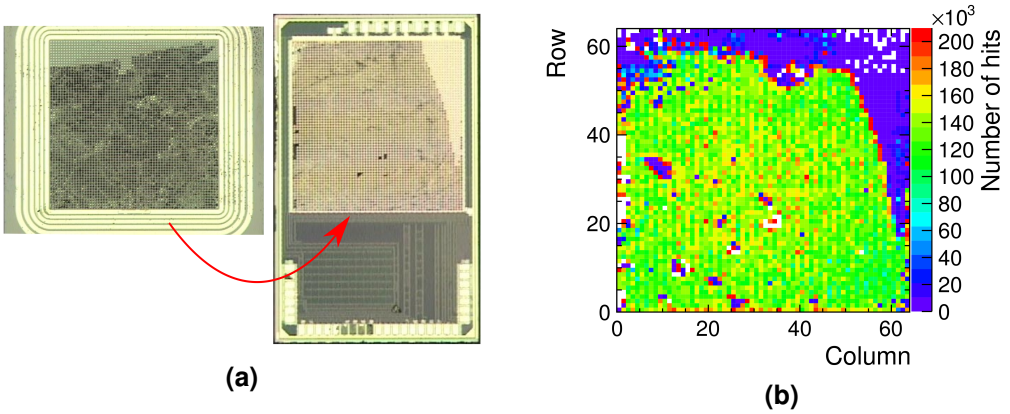


Figure 6.8: (a) Images of the sensor and CLICpix with the bumps applied just before the completion of the flip-chip of assembly 43, indicated by the red arrow, and (b) the corresponding ^{90}Sr-hitmap.

such as missing bonds but cannot be distinguished with such pictures.

A further defect that could arise from the bump-bonding process is, if one or more pixels are connected together through the solder spreading to the neighbours causing a short. A hitmap would not be able to easily provide this information as the exact location of the particles' hit position is not known. Instead, the quality of the bumps is determined using test pulses that are built in to each pixel on the readout chip and injected into the CSA. The matrix was split into sixteen 4×4 blocks. One pixel in each block was pulsed for a thousand frames and the ToT response of the pixels was measured. All of the nearest neighbours of the pulsed pixel (within ±1 pixel in the row and/or column direction) were checked and if in each frame it had a ToT above 0, then it was considered to be connected to the neighbour. This check was done for all pixels in the matrix and the resulting plots are shown in figure 6.9. The percentages of pixels that are considered as non-responsive (dead), weakly responsive and shorted are shown in table 6.3 for each assembly.

Table 6.3: The quality of the pixels for each assembly.

Assembly	Dead pixels	Weakly responding	Shorted pixels	Good pixels
14	0.2%	11.1%	1.7%	87.0%
31	0.3%	1.8%	1.0%	96.9%
43	3.0%	22.5%	1.0%	73.5%
50-4	13.4%	39.2%	1.7%	45.7%

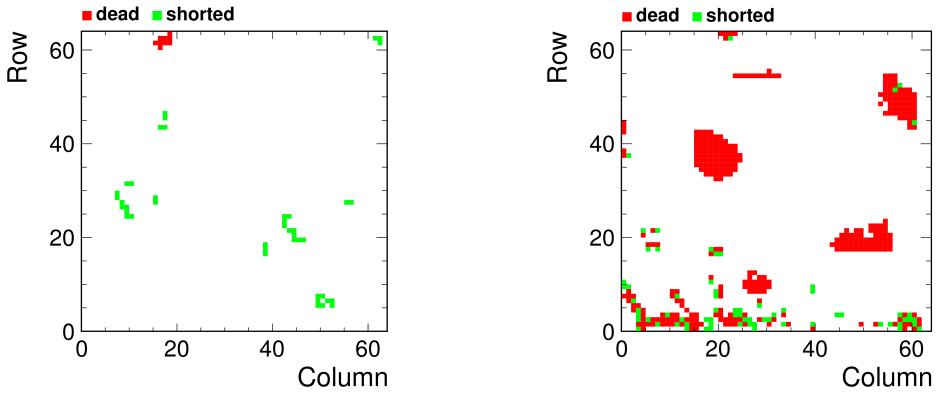


Figure 6.9: The location of shorted pixels (green) and dead to the test pulses (red) for two planar assemblies: (a) 31 and (b) 50-4.

6.2.3 Calibration of ToT and energy

The time-over-threshold measurement depends on the type of chip and the settings used, such as the clock frequency, the time it takes the discriminator to return to baseline and the threshold value. As a result, the ToT measurement varies from assembly to assembly. Therefore, to be able to compare the measurements across all the assemblies, a calibration is needed. This converts the ToT into the physical unit of energy deposited in the sensor or the number of electron-hole pairs collected. To achieve this, the ToT response of the assembly is observed using sources of known energy to produce the signal. This can be done either by test pulse injection or using characteristic X-ray photons from a radioactive source or a fluorescence spectrum.

Test pulse calibration

The first type of calibration carried out was the test pulse calibration which aims to provide a calibration between the ToT measurement of the readout chip and the collected charge. The calibration takes advantage of the test pulse injector built into each pixel on the readout chip. A capacitor is charged up to a programmable voltage, then discharged to produce a signal of known charge into the front-end amplifier in the readout chip. The amount of charge depends on the voltage and on the value of the capacitor, which is designed to be 10 fF with a variation of $\pm 20\%$ given by the foundry. The relationship between the pulse height (P) and the ToT is parametrised by the following surrogate function:

$$\mathrm{ToT} = a\mathrm{P} + b - \frac{c}{\mathrm{P} - t}, \tag{6.1}$$

where a, b, c and t are the fit parameters [82]. This function consists of two parts: a

linear part and a non-linear part. The linear term $aP + b$ becomes dominant for larger pulse heights with gradient a and intercept b. On the other hand, the non-linear part $c/(P - t)$ dominates for low pulse heights with the curvature described by c and an asymptote at t. Another important feature of this curve is the point at which it crosses the x-axis, i.e. has a ToT of 0. This defines the threshold at which a signal can be detected.

Similar to the connection quality test, the matrix is split into 4×4 blocks and one pixel in each block is pulsed 100 times. This time the non-pulsed pixels were masked. The voltage of the test pulse is varied from 2–400 mV in steps of 2 mV, and the mean ToT response for the 100 pulses is calculated. The results have been separated into even and odd columns, shown in figure 6.10, due to a known bug in the CLICpix circuity which creates an asymmetry between the columns. The asymmetry occurs because the CLICpix has a double column structure and there is slightly different layout of pixels on the left and right side of the double column [53]. This different layout causes a constant charge injection in the discriminator of odd columns, resulting in odd columns having a slightly lower threshold value, and thus having higher ToT values. This difference, illustrated in figure 6.11, results in a faster turn on for the odd columns producing a much steeper initial rise in the non-linear region. However, the linear regions have a similar gradient. This results in the b and c fit parameters changing systematically between columns, while a is roughly the same, as will be demonstrated later.

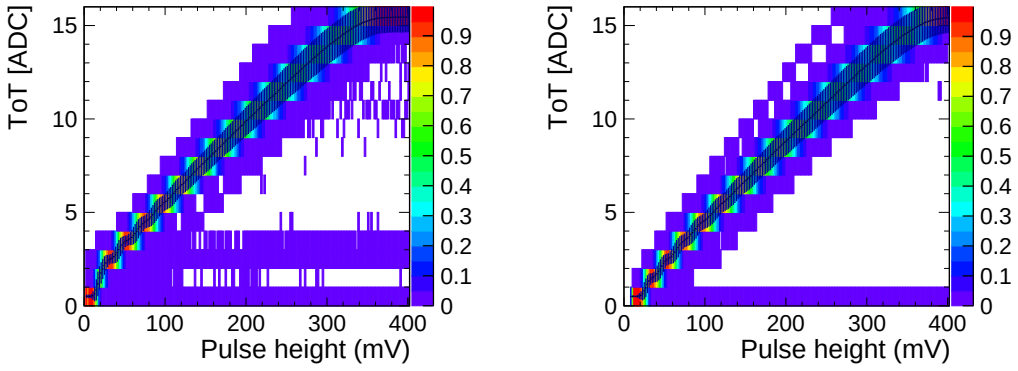


Figure 6.10: The ToT response of assembly 31 as a function of test pulse height split into (a) odd and (b) even columns. The black points show the mean and RMS of the 2D histogram.

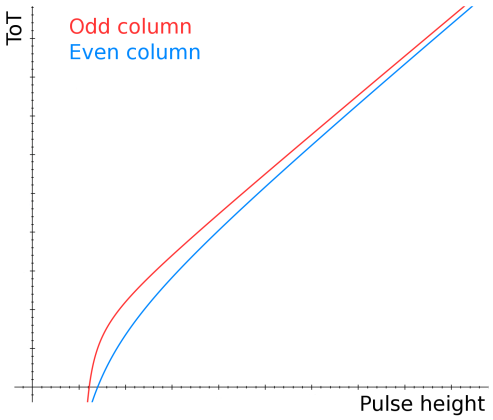
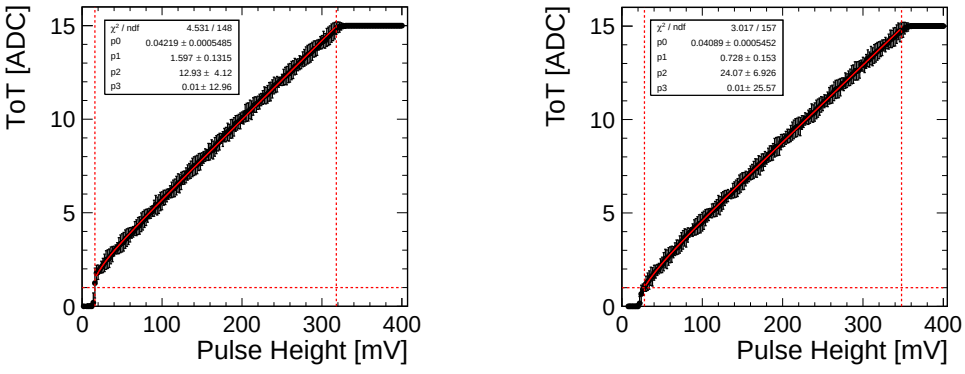


Figure 6.11: An illustration showing the difference between odd and even columns' surrogate function response.

For each pixel, a curve was produced and fitted with equation 6.1, examples for odd and even columns are shown in figure 6.12. To generate a better fit, a few constraints were put in place. The first one was to limit the range of the pulse heights fitted. The lower limit was put in place to reduce the number of unresponsive pulse heights fit and the higher limit was used so that the pulse heights with a saturated ToT were not fit. These limits vary according to each pixel response and are illustrated by the dashed red lines in figure 6.12. The lower limit is found by taking the pulse height at which the mean ToT is above 0.5 then subtracts 5 mV from this. The second constraint was to limit the values which the fit parameters could take.



(a) **(b)**

Figure 6.12: The ToT response of assembly 31 as a function of test pulse height for two select pixels: (a) (33, 35) and (b) (60, 48). The black dots are the mean values with the error bars indicating the RMS. The bold red line indicates the fitted surrogate function within the limits given by the dashed red lines.

The values of the fit parameters for each pixel are shown in figure 6.13 for assembly 31 and figure 6.15 for assembly 43. Here the column structure previously mentioned can be clearly seen. These fit parameter values for each pixel can be used to calibrate the beam test data into physical units. This is done by inputting the measured ToT value into equation 6.1 to obtain a voltage and converting that into the number of electrons. This also produces a more uniform response across the matrix by accounting for the pixel to pixel variation. To study the amount of variation in the fitted parameters, a histogram was plotted (figure 6.14 for assembly 31 and figure 6.16 for assembly 43) for each set of columns showing the mean and RMS. For both assemblies there is a periodic structure in the t parameter histogram for the odd columns. Since this only occurs in the odd columns it is assumed that this is due to the known bug in the CLICpix. The large number of entries seen in the b parameter plot for even columns of assembly 43, figure 6.16(b), is due to the dead pixels, the high threshold value of the assembly and the smaller curvature of the calibration curve (low c value) for even columns.

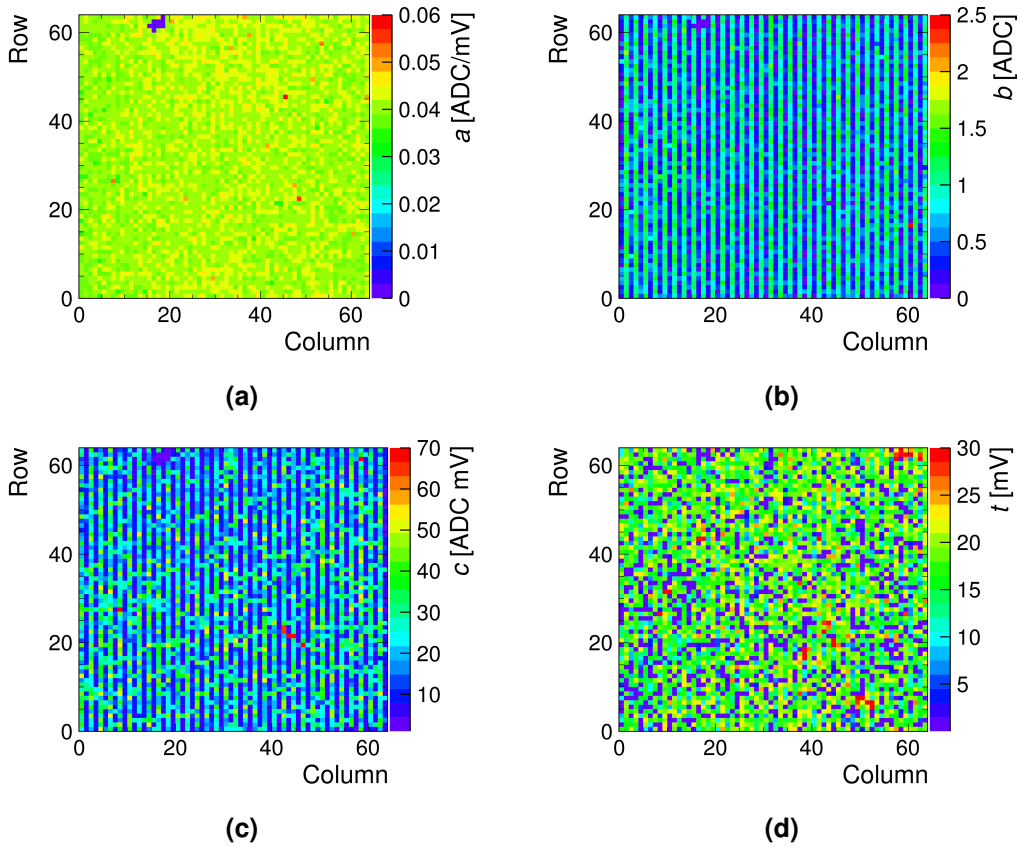


Figure 6.13: The fit parameters of the surrogate function (equation 6.1) across the matrix for assembly 31: (a) a, (b) b, (c) c and (d) t. Each fit has been split into odd and even columns due to the known issue of the CLICpix.

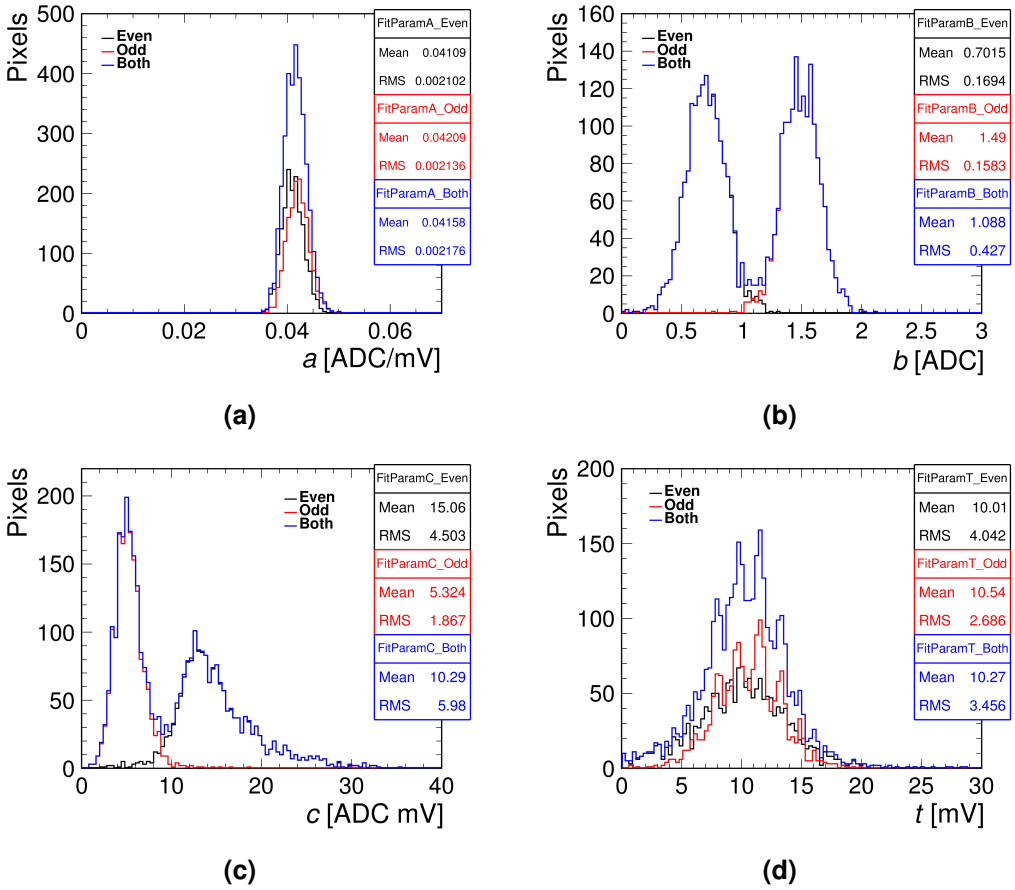


Figure 6.14: The fit parameters of the surrogate function (equation 6.1) in a histogram for assembly 31: (a) a, (b) b, (c) c and (d) t.

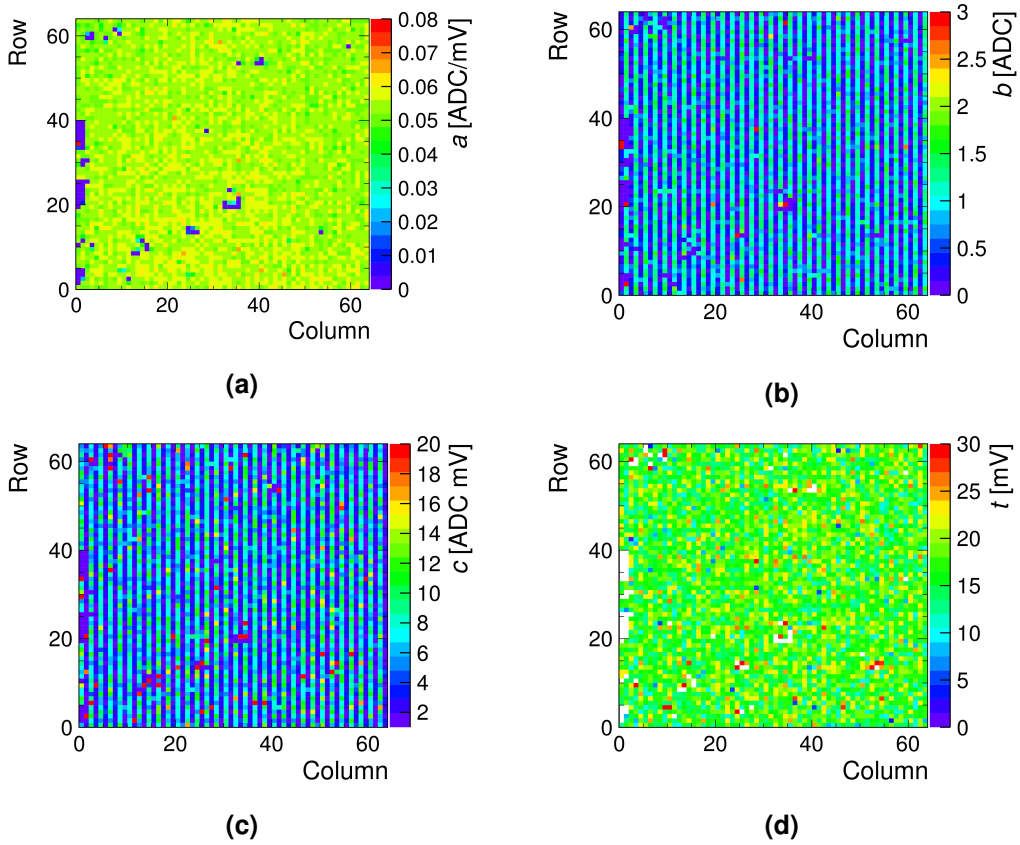


Figure 6.15: The fit parameters of the surrogate function (equation 6.1) across the matrix for assembly 43: (a) a, (b) b, (c) c and (d) t.

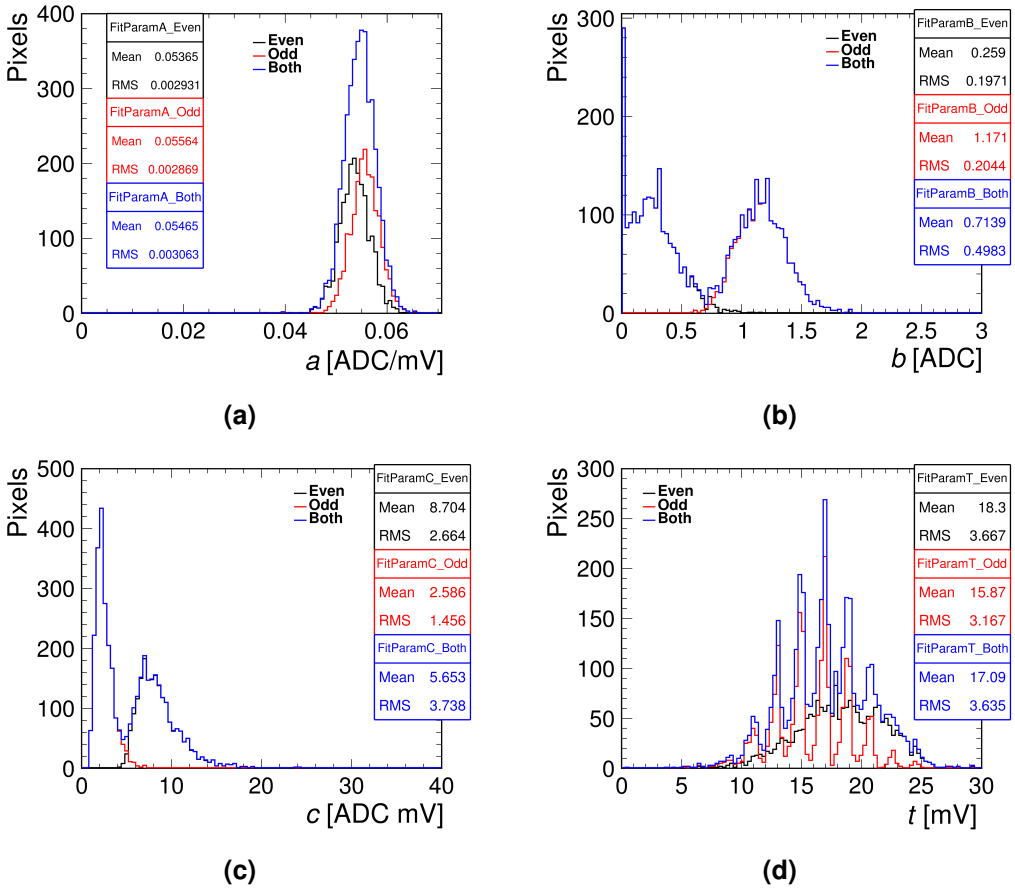


Figure 6.16: The fit parameters of the surrogate function (equation 6.1) in a histogram for assembly 43: (a) a, (b) b, (c) c and (d) t.

Source measurements

Using the test pulse calibrations to convert the measured ToT into physical units relies on accurate knowledge of the test pulse capacitor. A method to verify this value is to use a ToT calibration by illuminating the assemblies with characteristic x-ray emission from a radioactive source. This takes advantage of the known energy peaks produced in the emission spectrum and the amount of energy required to generate an electron-hole pair in silicon, the ionisation energy, 3.6 eV. Utilising this knowledge and the results from the test pulse calibration, the ToT response can be converted to an absolute energy value to allow a comparison between the measured peak values and the known peak energy.

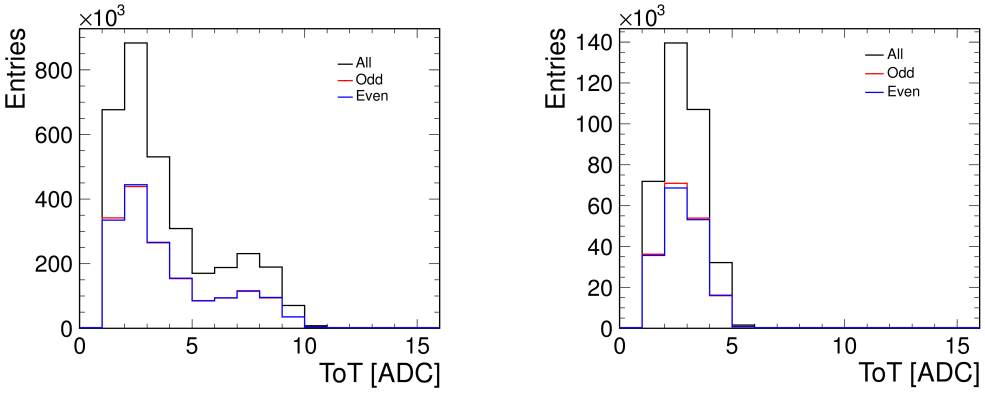
All the analysed data in the source measurements are restricted to single hit clusters. As a result, the deposited energy is not shared over several pixels and an accurate measurement of the energy can be made. However, when there is charge sharing to neighbours, but the energy collected is not large enough to cross the threshold,

the measurement would be below the actual energy deposited. Another undesirable feature is when pixels receive hits too many times due to noise. Pixels were deemed to be noisy if they were more than 3 standard deviations above the mean rate. These values were obtained from a Gaussian fit to a histogram of the number of times each pixel received a single hit cluster. In total around 3–10%, depending on the source and assembly, of pixels were noisy and subsequently masked. The data for each assembly was looked at for the whole matrix and not on a pixel-by-pixel basis due to the low number of hits for the single pixels.

This type of calibration was performed with two sources: ^{241}Am and ^{109}Cd. The reason for not using the ^{90}Sr source is that the peak energy is too high and would saturate the ToT counter of the CLICpix. The used sources produce peaks in the measured response and their spectra are shown in figure 6.17. The ^{241}Am emission X-ray spectrum has two peaks, one at 59.5 keV and one at 26.3 keV. While for the ^{109}Cd, there is one peak at 22.1 keV. In both cases the peaks can be seen but are at the lower end of the ToT spectrum. This is a problem as this region is potentially close to the noise and is in the non-linear part of the surrogate function. The largest energy peak from the ^{241}Am is not very well defined but still visible. Splitting the results into even and odd columns made no impact on the values of the peaks, suggesting that the resolution of the ToT counter is not fine enough to resolve this difference.

Results from the test pulse calibration can be used in conjunction with the measured ToT spectrum from the source calibration to estimate a value for the test pulse capacitor. This is done by obtaining the pulse height value of the ToT peaks seen in the source spectra from figures 6.10(a) or 6.10(b). This is first converted to a charge using: $Q = CV$ (using the nominal value of the test pulse capacitor), and then converted to electrons by dividing by the elementary charge of an electron. This is then multiplied by 3.6 eV (the energy it takes to generate an electron-hole pair in silicon) to produce a final value of the measured deposited energy. A ratio is then taken between the known peak value and the measured peak value. This is the factor by which the sensor test pulse capacitance needs to change, which is added to the nominal capacitance value to get the measured capacitance value. For assembly 31, this results in a value of around (18.1 ± 2.2) fF for odd, and (15.5 ± 2.5) fF for even columns respectively. Both these values are larger than the foundry value of 10 fF and not within the expected error of 20%, indicating a problem with the source measurements.

The coarse nature of the ToT spectrum means that the peaks cannot be properly resolved and one ToT value can cover a large range of energies. This is evident by the lower ^{241}Am peak and the ^{109}Cd peak having the same peak ToT value of 2, while having a difference of 4.2 keV in their known values. This means that the value for the



(a) ^{241}Am **(b)** ^{109}Cd

Figure 6.17: The ToT spectrum from the radioactive sources of (a) ^{241}Am, with peaks at ToT 2 and 7, and (b) ^{109}Cd, with a peak at ToT 2, for assembly 31.

test pulse capacitor cannot be determined accurately and is subject to large systematic uncertainties. One method to overcome this would be to look at the spectrum produced pixel-by-pixel. However, this was not possible due to the low number of statistics available. Another option would be to tune the settings on the CLICpix in order to better suit the measurements but this was not possible as the CLICpix setting would no longer match those in the beam tests. As a result, a charge to ToT calibration could not be performed using radioactive sources, instead the ones taken with the test-pulse injection were used.

6.2.4 Threshold DAC calibration

The global threshold is controlled by the 12-bit DAC in the periphery of the readout chip. These DAC values need to be converted into physical units to estimate the minimal detectable signal and to allow for a comparison between chips. In a similar method to the ToT calibration, the procedure is to inject a known voltage into the CSA using the test pulse generator in each pixel cell on the readout chip. A scan through the threshold DAC values is done to produce a turn on curve (S-curve) for each pixel, from which the noise and threshold of the pixel can be extracted. This is done for several pulse amplitudes, allowing for a correlation between the input charge and the threshold DAC value. This was done only with assembly 31.

The method implemented in this thesis was to separate the pixels into 4x4 blocks and pulse one pixel in each block simultaneously with the remaining pixels in the block masked. Therefore, for each mask stage there would be 256 pixels pulsed at the same time. At each threshold value the pixels were pulsed 100 times and the number of hits seen on each pulsed pixel was recorded. As the threshold decreases (an increase

in threshold DAC values due to the polarity of the chip, collecting electrons), the pixel registers a hit more often, until the pixel sees a hit for all 100 pulse injections. The shape of the resultant curve is that of an elongated S, hence the name S-curve. Once this is done for all pixels, the curves are fit with the following sigmoid function:

$$\text{Counts} = \frac{h}{2}\left(1 + \text{erf}\left[\frac{x-T}{w\sqrt{2}}\right]\right), \tag{6.2}$$

where h is the height, T is the point at which half the height is reached (threshold) and w is the width of the S-curve (noise RMS). An illustration of the fit and its variables is shown in figure 6.18 for one of the pixels of assembly 31. The values of the fitted parameters for assembly 31 with a test pulse voltage of 0.03 V are shown in figure 6.19. The column structure as well as the dead pixels are clearly visible. A summary of the fit parameters for the different column types are shown in table 6.4.

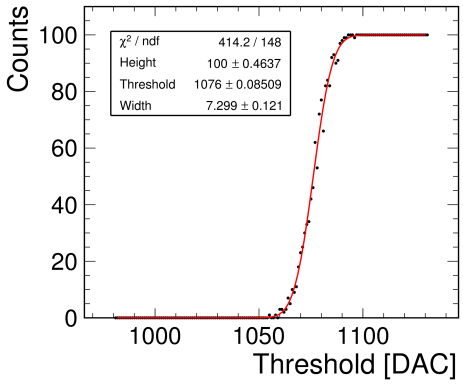


Figure 6.18: An example of a fitted S-curve for one of the pixels of assembly 31 for an injected pulse voltage of 0.03 V.

Table 6.4: The mean and RMS of the S-curve fit parameters at a test pulse amplitude of 0.03 V.

Assembly	Column type	Threshold [DAC]		Width [DAC]		Height [counts]	
		Mean	RMS	Mean	RMS	Mean	RMS
31	all	1049	30.1	7.16	1.62	99.55	0.71
	odd	1024	17.0	5.73	0.71	99.97	0.05
	even	1074	17.1	8.63	0.74	99.12	0.80

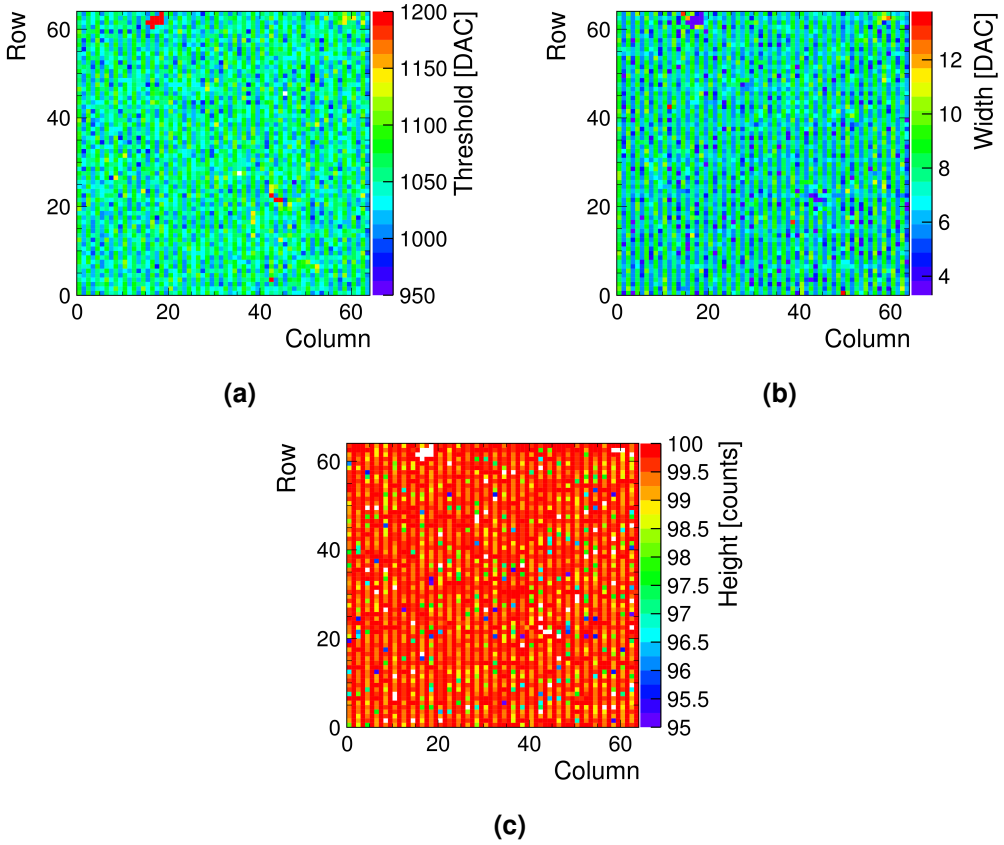


Figure 6.19: The fitted S-curve parameters across the matrix for assembly 31 at a test-pulse amplitude of 0.03 V, showing the (a) threshold, (b) width and (c) height.

Before the DAC calibration is done, several selection cuts are made to ensure that no bad fit pixels skewed the final results. These cuts were performed for each test-pulse amplitude value. The first cut was based on a histogram of the χ^2/NDF of the fits, shown in figure 6.20. Pixels were removed if their S-curve fit χ^2/NDF value was 3 standard deviations above the x-value of the highest bin of the histogram. Another cut was to remove pixels for which the fitted height was below 95, which had a width greater than 20 or for which the fitted threshold was not within the threshold scan range, this range depended on the test-pulse amplitude used. For all test-pulse amplitude values used, fewer than 5% of pixels were cut from the analysis.

After these cuts, the mean of the fitted thresholds is plotted as a function of the injected charge which has been converted from the voltage of the injected pulse. This is done using the nominal value for the test pulse capacitance, 10 fF, and the resulting plot is shown in figure 6.21. The gain of the threshold DAC is the reciprocal of the fitted gradient. Each DAC step corresponds to around $(11.2\pm0.4)\,e^-$ for the odd columns, $(12.2\pm0.4)\,e^-$ for the even columns and $(11.7\pm0.5)\,e^-$ for both columns. These are

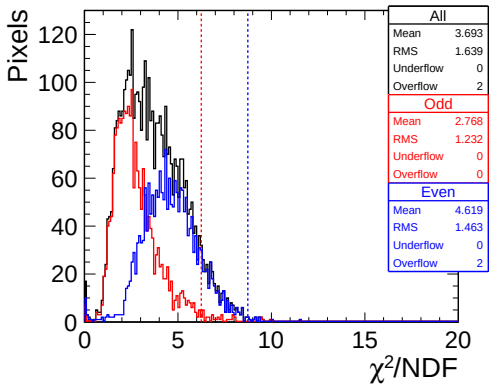


Figure 6.20: The χ^2/NDF values of the S-curve fits for assembly 31 at a test-pulse amplitude of 0.03 V. The vertical lines indicate the value of the cut points for both even (blue) and odd (red) columns.

close to the design value of $\approx 10\,e^-$ per DAC step. From figure 6.21 which is fit with a straight line: $y = ax + b$, the operation threshold and noise RMS of assembly 31, shown in table 6.1, can be given in electrons. these are $1137\,e^-$ and $89.2\,e^-$, respectively.

6.3 Capacitively coupled assemblies

In addition to planar assemblies, measurements were also performed on the capacitively coupled pixel detectors, described in chapter 4.3.3. Various lab measurements have already been shown in [52, 83]. An important factor in this type of sensor is the glue layer and how this impacts the coupling between the two chips. This is because when the signal is collected and processed by the CCPDv3 it then transfers the signal

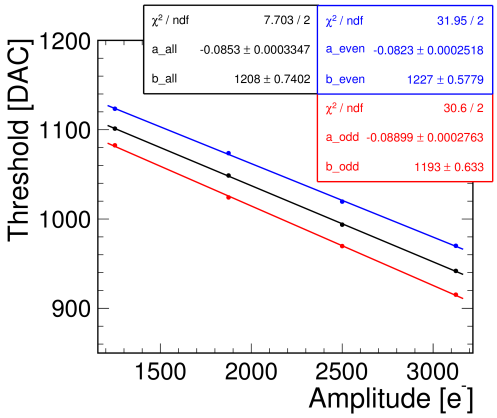


Figure 6.21: The mean over the matrix of the fitted threshold as a function of input charge for assembly 31. Also shown are the results of the separate linear fits for even (blue), odd (red) and all pixels (black).

to the readout chip capacitively via a thin layer of glue.

This thesis will focus on the ToT calibrations which aim to characterise the properties of the sensor, glue and readout chip. This was done using a radioactive source and is used in the conversion of simulation results to ToT. This allows for a comparison of simulation and measurement for the beam test results detailed in chapter 7. The bias voltage at which these measurements were taken was –60 V at a setting threshold of 1020 DAC, which match those used for the beam tests. The beam test results are for one sensor, SET 13, and will be discussed in chapter 7. So the only calibration results shown are those of SET 13.

The calibrations were conducted in the lab using a ^{90}Sr source of activity 29.6 MBq, providing photons that have a peak at 196 keV in their energy spectrum. The source was shone onto the back of the CCPDv3 and the analogue response of the 2nd stage amplifier within a single pixel of the CCPDv3 was monitored with a fast sampling oscilloscope. This produces a value of the amplifier voltage as a function of time. To obtain the pulse height from this, the baseline voltage of the amplifier was subtracted then the peak was inverted. After this, a peak of the pulse height is defined as the region where the voltage exceeded 90% of the raw pulse height given by the oscilloscope. This final pulse height value is then the amplitude of the Gaussian fit across this region. An example of a peak fit is given in figure 6.22. This analysis code was developed by S. Green and details can be found in [83].

During the data taking the CLICpix shutter would be left open, waiting for a trigger from the oscilloscope, after which it would read out the hits. In order to ensure that the charge measured corresponds to a single hit, only frames with a hit counter value of one are used. This is done for a select number of pixels along the first row, 16 in total, as these have the circuitry necessary to be able to read out their analogue signal.

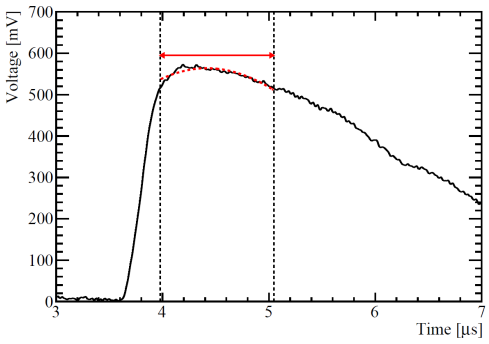


Figure 6.22: An example of the fitted pulse height [83]. The vertical dashed lines represent the fit region and the dashed red line is the Gaussian fit.

Only 14 pixels were used in the analysis as two showed an unusual response. The measurements were performed one pixel at a time for all pixels. After 12k frames had been recorded the next pixel was measured. A threshold was set on the oscilloscope to be 20 mV. This was chosen to be above the noise while still being able to calibrate the low pulse height values which are of importance for the charge sharing hits in the simulation. All the pixels are in odd columns meaning that the data could not be split into even and odd to accommodate for the CLICpix charge injection issue. Once all the data had been collected for each of the 14 pixels, it was combined together.

By comparing the analogue output of the CCPDv3 to the ToT measurement on the CLICpix, a calibration of the assembly can be made. The resulting pulse heights on the CCPDv3 and ToT values of single hits of a ^{90}Sr source are shown in figure 6.23 for unfiltered data prior to cuts being applied. From the pulse height spectrum, a peak is hard to distinguish and the amplifier saturation can be seen with no events above 750 mV. For the ToT spectrum, a clear peak can be seen at a ToT value of 5 and the saturation of the ToT counter can also be clearly seen by the large number of events with a ToT value of 15.

The ToT as a function of pulse height is shown in figure 6.24(a). At low pulse heights the curve is susceptible to noise in the pulse height, as demonstrated by the entries with low ToT values below 50 mV. A potential source of these noise hits, is that low pulse heights are also harder to distinguish from the baseline voltage jitter making them harder to fit correctly. However, this does not explain the large ToT values, > 8, seen at low pulse heights. These could possibly be due to a pulse triggering the oscilloscope but not the CLICpix, then another hit arrives triggering the CLICpix to read out and have a large ToT for a small pulse height. At higher pulse heights, mismatched events are

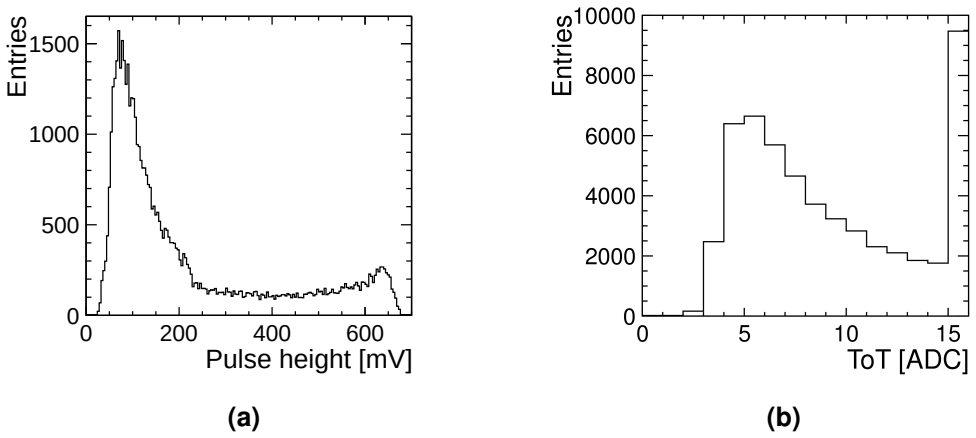


Figure 6.23: (a) The pulse height and (b) ToT spectrum measured with the ^{90}Sr source.

found around a ToT of 4 which is the peak region in the ToT spectrum. In addition to this, there is a lack of data points in the central part of the curve between 250–450 mV. This reduction in data is due to the spectrum of the ^{90}Sr source not having as many events in this energy region. The saturation of the ToT counter, a ToT value of 15, is also seen at pulse heights above 350 mV.

To obtain the calibration, a projection along the x-axis was done to produce a histogram of the ToT every 4 mV. The histogram was then fitted with a Gaussian curve and the mean extracted. So as to not obtain ToT values outside the valid range, several checks were made. The first check was to see if the mean of the Gaussian was less than 0 and greater than 16. The second check was to see if the RMS of the histogram was 0. If both of these were true, the mean from the histogram was used instead. This procedure produces the plot shown in figure 6.24(b). This was chosen as opposed to using an MPV value as it produced a smooth curve rather than a step function. From the plot, the effect of the noise and large ToT values at low pulse height can be seen, causing a rise in the points at low pulse heights.

So that this may be used as a conversion of pulse height to ToT in the simulation, the calibration curve was fit with the same surrogate function given in equation 6.1. The low number of data points in the middle, along with the saturation of the ToT, cause problems for the fitting function as the linear part is now split into two regions. This meant that a satisfactory fit was not possible with just one surrogate function. To address this, the data was fit in two parts one for the low range data: 0–200 mV and one for the high range data: 200–620 mV. This is an unphysical fit but as the simulation produces pulse heights below 200 mV, only the lower calibration curve was used. The

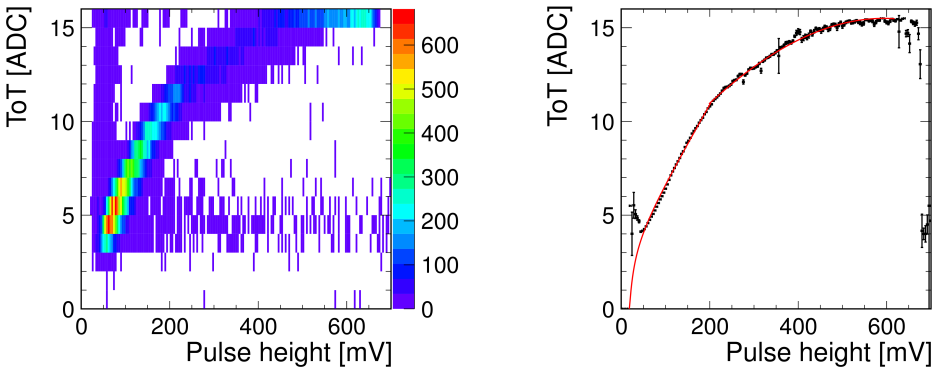


Figure 6.24: (a) The ToT response as a function of the pulse height from the CCPDv3 for SET 13 for the data before the selection cuts. (b) The mean of the Gaussian fits for the final data with the surrogate fits overlaid (red).

nature of the curves is very dissimilar to those found in figure 6.12 given by the test pulse, where very distinctive non-linear and linear parts are seen. The reason for this is that there is a potential for a wrong association of events due to the CLICpix waiting for the trigger from the oscilloscope. Another reason is that the method for obtaining the pulse height sometimes fails to fit the peak properly due to distorted pulse shapes seen in the amplifier.

6.4 Summary

Calibrations and characterisation of planar sensors bump-bonded to the CLICpix readout ASIC have been carried out. Initial current-voltage characteristics for all assemblies were conducted, showing their breakdown voltages to be above the full depletion voltage and low current values. In determining the quality of the bump bonds, assembly 31 had the largest number of good pixels, 97%. However, the number of pixels that were either dead, weakly responding or shorted ranges from 3–54%. The reason for these defects are due to the bump-bonding process and they could be seen before the sensor and readout chip were flipped and joined together. This low yield value was not surprising due to the experimental nature of the bonding process at small pitches of 25 µm on the chip-to-chip level.

The test pulse calibration provided a method to calibrate the ToT into physical units with more reliable results than the source measurements. The drawback is that the value of the test pulse capacitor needs to be known for each pixel. To obtain this, source measurements were attempted, but reliable results could not be obtained. This is because of the limited precision of the ToT counter, 4-bit, not allowing the peaks of the radioactive sources to be resolved with enough resolution. This factor has been taken into account and the next generation of chip, the CLICpix2 has a 5-bit ToT counter. The threshold DAC calibrations yielded a conversion of (11.2-12.2) e^- per DAC step in agreement with the design value of 10 e^- within 20%. This indicates that the test-pulse capacitance is close to the nominal value, supporting the conclusion that the source ToT calibrations are not reliable is justified.

In all measurements the analysis of data was two-fold because of the odd even column asymmetry. Therefore, it was harder to interpret the results or verify the uniformity of the chip. Again, this problem has been addressed in the CLICpix2 which is currently under test.

In the ToT calibration of the capacitively coupled assemblies, a ^{90}Sr radioactive source, which emits photons of known energy, was used. The analogue output of the CCPDv3 sensor was compared to the ToT response of the CLICpix. The surrogate function

(equation 6.1) was unable to fit the data due to the nature of the source's spectrum and the saturation of the ToT. In addition, low ToT and pulse height values could not be distinguished from the noise. These problems will be resolved with CLICpix2 and by using an x-ray source to better control the energy of the incident photons, allowing to produce enough data points over the full ToT range.

CHAPTER 7

Beam tests of capacitively coupled HV-CMOS sensors

One way to test the performance of the sensor and readout chip is to use beam tests. This is where a high energy beam is sent through the device to extract its measurable properties, such as resolution and efficiency. Beam tests make up an important part of detector R&D, allowing for an evaluation of the prototype sensor. Often the results will be used, in conjunction with laboratory characterisations and simulations, to inform the design of new chips or examine the validity of a sensor design for full-scale production.

This chapter sets out to describe the conditions of the test-beam and the set-up used for taking measurements with capacitively coupled HV-CMOS sensors. Once the data is acquired, the offline analysis will be used to generate the results. A description of the techniques used will be given, these include: alignment, tracking and cluster algorithms. After this, the results will be shown with comparisons to simulations, where applicable.

7.1 Experimental setup

A common setup for test-beam measurements is to have the device under test (DUT) at the centre of a so-called telescope. The concept of a telescope is to have several sensors on either side, which are called the telescope planes, to provide information that is used to reconstruct the tracks of the charged particles. Using this, along with the data from the DUT, the tracking performance of the sensor can be obtained.

7.1.1 Test-beam set-up

The CERN SPS provides a secondary beam of charged pions and protons with a momentum of 120 GeV/c to the beam area on the H6 beam line located at the CERN North Area. The beam spills last for 4.8 s producing around 6–9 $\times 10^5$ particles and typically occur at intervals of 20–40 s.

7.1.2 Device under test

In this thesis only one DUT was studied, named SET13. This is a CCPDv3 HV-CMOS sensor capacitively coupled to the CLICpix readout chip in a so called ideal alignment, illustrated in figure 7.1. This means that the centre of gravity of both metal pads (including the metal via on the CLICpix) were aligned in order to minimise the cross-coupling

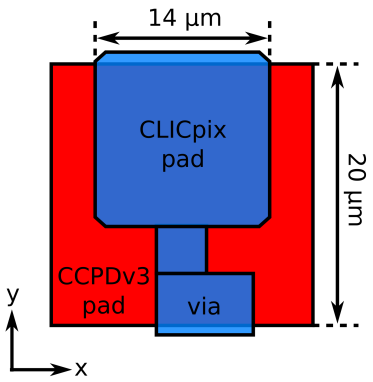


Figure 7.1: A sketch of the so called ideal alignment of the CCPDv3 pad, in red, and the CLICpix pad, in blue.

to neighbours, maximising uniformity in the matrix. The DUT was operated at a substrate bias of –60 V, the foundry recommendation. The chip was calibrated by adjusting the 4-bit threshold on each pixel to correct for the threshold dispersion. From this, the baseline was found to be at a DAC (digital-to-analogue converter) value of 1140. By design, one DAC step corresponds to approximately 10 e^- [52]. An operating threshold DAC value of 1020, which is equivalent to around 1000 e^-, was set so that the threshold was above the noise level[1]. At this threshold around 70 noisy pixels were found (less than 2% of all pixels) and are subsequently masked. These are found by fitting a Gaussian to a histogram of the number of hits for each pixel. Then the noisy pixels are defined as pixels which are more than 5 standard deviations above the mean rate of response. The pixel mask used for the sensor also masked the last four columns, as these only have the first stage amplifier in the sensor. The pixels that are adjacent to the edge of the matrix were masked as charge could flow to the edges of the chip and be not be collected, spoiling the measurements. In addition, a circular mask was applied to the matrix in order to account for the variations in the coupling strength due to fabrication; discussed further in chapter 7.4.1. A clock frequency of 20 MHz was set for the CLICpix time-over-threshold (ToT) measurements, with a constant discharge current, as described in chapter 6.1.

7.1.3 AIDA beam telescope

The AIDA beam telescope [84] consists of two mechanical arms which house three telescope planes each and which are mounted in aluminium jigs providing accurate positioning, as shown in figure 7.2. They are installed on rails which allow the planes to be positioned with a minimum distance of 20 mm between them. The sensors used for the planes are the MIMOSA 26 [85] monolithic active pixels sensors (MAPS)

[1] A decrease in DAC values means an increase in threshold, due to the negative polarity of the chip.

that provide precise spatial measurements of the incoming particles. They consist of 18.4 µm×18.4 µm square pixels organised into 1152 columns and 576 rows, creating an active area of 21.1 mm×10.6 mm of silicon. The planes produce a pointing resolution at the DUT of $\sim$ 1.6 µm [52]. The DUT is placed in between the two mechanical arms at the centre of the telescope and is mounted onto a precision rotation stage. The measured position along the beam line of all the sensors in the telescope are given in table 7.1. The mechanical arms sit on top of a table that is split into two parts to allow adjustments in the position of the system to be made with respect to the beam axis. The coordinate system used for the telescope is a right handed one where the x-axis is perpendicular to the beam and parallel to the floor, the y-axis is perpendicular to the beam and perpendicular to the floor pointing down and the z-axis points along the beam line. These define the global coordinates. The local coordinates are the ones that describe the position on the sensors.

The telescope active area is much larger than that of the DUT, 1.6 mm×1.6 mm, therefore to improve the efficiency of the data taking, the CLICpix shutter was held open for several of the MIMOSA's 230 µs rolling-shutter periods. After this period, the frame on the CLICpix was attached to the several frames that occurred on the telescope during this time. In addition to this, a FE-I4 [86] trigger plane was used in conjunction with three crossed scintillators at upstream and downstream of the telescope, to select tracks in a region of interest (RoI), taking advantage of the FE-I4's functionality (ability to select certain areas on the chip), as seen in figure 7.3.

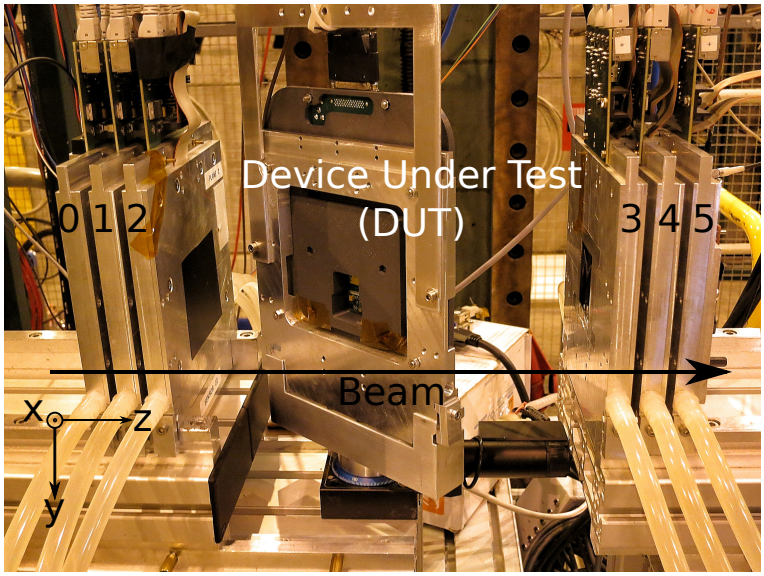


Figure 7.2: The AIDA beam telescope in the H6 beam line at the CERN SPS North Area, showing the six telescope planes and the device under test mounted onto a rotation stage. It also defines the coordinates of the system.

Table 7.1: The sensors in the telescope and their positions along the beam line (z-axis).

Detector ID	z [mm]
Mimosa 0	0
Mimosa 1	20
Mimosa 2	40
DUT	153.5
Mimosa 3	321
Mimosa 4	341
Mimosa 5	362

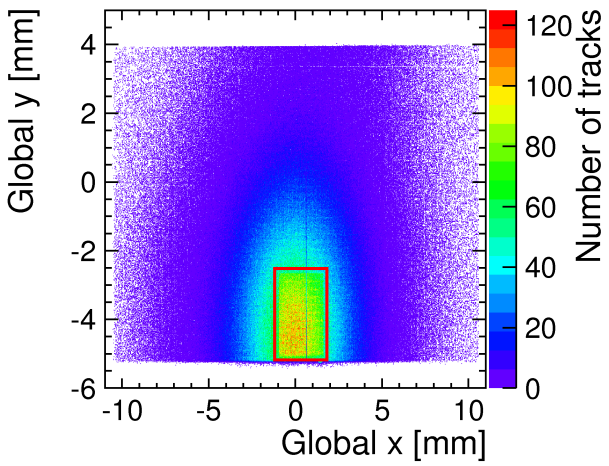


Figure 7.3: The track intercept positions on one of the MIMOSA 26 sensors in a telescope plane, highlighting, in red, the region of interest selected on the trigger plane.

7.2 Reconstruction method

The analysis of the test-beam data requires several reconstruction steps to produce the final results. Depending on the sensors used for the telescope and the DUT, the reconstruction algorithms may need to vary between the two, e.g. the MIMOSA 26 sensors have use binary information compared to the digital information (4-bit ToT counter) of the CLICpix. The following sections will focus on the algorithms used for the DUT.

7.2.1 Clustering

The first piece of information that is processed are the hits on the sensor. A hit is deemed to have occurred when a pixel measures a charge that is above the threshold of the pixel within the shutter of the readout chip. The hits are used to create clusters. A cluster consists of adjacent hit pixels and is typically surrounded by empty pixels and/or it is alongside a sensor edge. For the sensors used in this study, this is the only criteria for a hit pixel to be part of a cluster, no timing information is used. The cluster algorithm

generates clusters by checking if a hit pixel has any hit neighbours, if this is true it is added to the cluster. This continues iteratively until there are no further adjacent hits to a cluster. The cluster is defined by two dimensions, the cluster x-width (the number of pixels within the cluster in the x-direction) and the cluster y-width (the number of pixels within the cluster in the y-direction). The x- and y-directions are the same as the ones defined in figure 7.2.

Each hit contains three components, the row and column number of the pixel, as well as the signal strength of the charge collected, recorded in ToT. To acquire the position of the cluster, a charge-weighted average of the individual pixel positions is used. This is called the centre of gravity (CoG). The position of the cluster with m hit pixels in the column direction, x, is given by:

$$x = \frac{\sum_{i=1}^{m} S_i x_i}{\sum_{i=1}^{m} S_i}, \tag{7.1}$$

where x_i is the x-position of a hit pixel in the cluster and S_i is its ToT value [87]. An equivalent procedure is also done in the y-direction. The cluster position can then be turned into a local position on the sensor and a global position in the telescope and used for analysis.

However, this does not always reproduce the track position effectively, in particular, when the width of the diffusion cloud is smaller than the pixel pitch. When a particle hits near the boundary between two pixels, a roughly linear charge sharing is seen, region B in figure 7.4; the width of this region depends on the width of the diffusion cloud. For hits outside this region, closer to the centre of one pixel (region A), the majority of the charge is collected by one pixel and the charge sharing is non-linear. As a result, the charge weighted position does not accurately reproduce the track position because this uses a linear charge interpolation. To address this, the so called eta (η) correction is used, to account for the non-linear charge sharing between neighbouring pixels [88]. This algorithm is suitable when charge sharing by diffusion is the dominant method of charge collection.

There are several ways to perform the eta correction. In this study two methods were used depending on the size of the cluster and will be described in subsequent paragraphs. The reason for using two methods was that neither method chosen would adequately cover all cluster size scenarios, a comparison between the two will be given later. To maximise the improvement to the resolution, the best method for a given cluster size was used. In both cases, the data needed to be analysed twice in order to apply the eta correction as knowledge of the eta values for all clusters was needed as well as the track position. The track position is the position on the DUT reconstructed by the tracking algorithm, more details will be given in the next section.

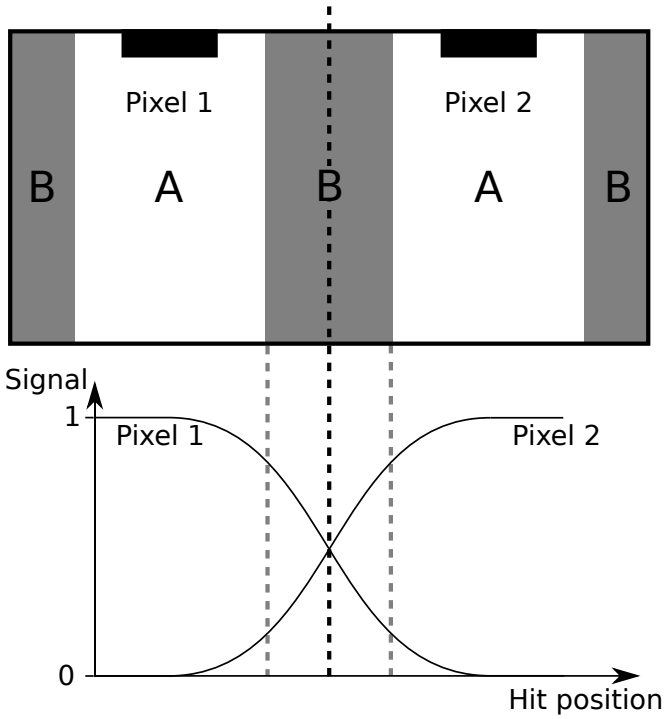


Figure 7.4: An illustration of the charge sharing between two pixels and the respective signal as a function of the hit position.

The first method, described in [89], was only used on clusters with a width of 2 in the x- or y-direction. For each cluster, the in-pixel position (the position within the pixel cell) obtained from the centre of gravity algorithm was plotted against the in-pixel track position, as illustrated in figure 7.5(a). This is done for all clusters after which a profile of the 2D histogram is taken and fitted with a fifth order polynomial, figure 7.5(b), in accordance with [89]. The data are then re-analysed and the correction is applied. This is done by putting the in-pixel CoG position into the fifth order polynomial function to obtain a corrected value. The result of this correction for all clusters of width 2 in the column direction is shown in figure 7.5(c) and the amount of correction this gives will be presented later in this section.

The second method, described in [87], is applied on clusters with widths of greater than 2 in x or y and implemented independently for widths of 3, 4, or 5 or more in x or y. It works out an eta value according to the equation:

$$\eta = \frac{S_R}{S_L + S_R}, \tag{7.2}$$

where $S_{L/R}$ is the ToT signal of the left or right pixel of the adjacent pair that has the largest signal in the cluster [87]. An eta value in the y- and x-direction is calculated.

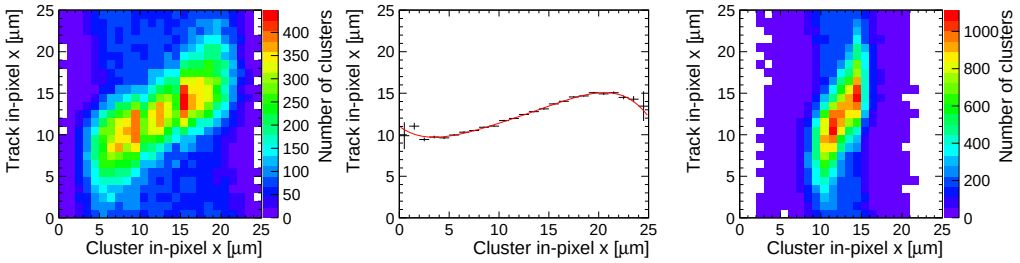


Figure 7.5: The in-pixel track intercept in the x-direction as a function of the in-pixel cluster centre-of-gravity x-position (a) before eta correction, (b) profile and polynomial fit (red) and (c) after eta correction for clusters of cluster x-width 2.

It corresponds to a weighted position of the track that gives the contribution (fraction) of charge that each pixel sees. For an eta value of 0 the left pixel collects all the charge, while for an eta value of 1 the right pixel collects all the charge. A typical eta distribution is shown in figure 7.6(a), while an example of an eta distribution in this study at perpendicular incidence is shown in figure 7.6(b). It can be seen that the example shown in figure 7.6(b) there are three peaks, two at the sides close to the centre and one peak at the centre. The shape of this distribution gives an indication of the level of charge sharing. The closer together the two peaks for each pixel are, i.e. further from 0 and 1, the larger the charge sharing and coupling capacitance (or cross-coupling, detailed in chapter 7.4.1) between adjacent pixels is [87]. Another influence the geometry of the sensor has on the eta correction is, the closer the pixel pitch is to the diffusion cloud length the less pronounced the peaks become. All these factors are the reason for the peaks, seen in figure 7.6, being difficult to distinguish. Due to the

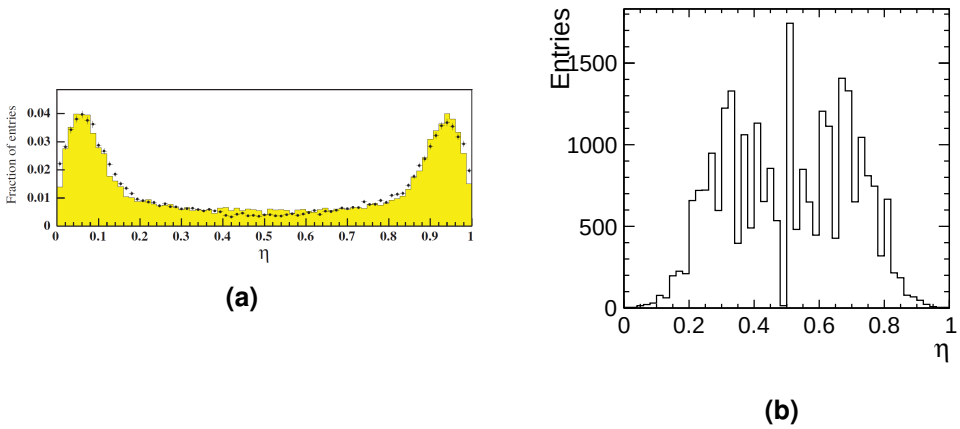


Figure 7.6: A typical eta distribution, from [87]. (b) An example of the eta distribution for clusters with cluster x-width 2 for this study at perpendicular incidence.

4-bit ToT counter there is also a limited number of values that the eta value can take. This also contributes to the peaks being difficult to distinguish. The large number of values in the bin at an η value of 0.5 may indicate that a considerable number of the pixels in the pair have the same ToT.

Once a distribution for each cluster size is found, the data is re-analysed. The corrected cluster position, x, is given by:

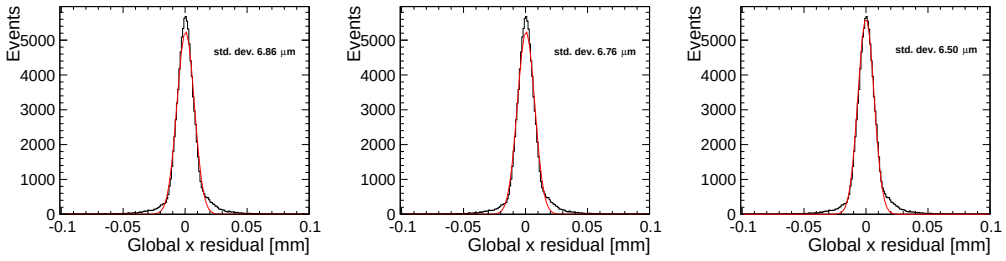
$$x = x_L + f(\eta)p\,, \tag{7.3}$$

where x_L is the position of the left pixel in the pair, p is the pixel pitch and the cumulative η function is given as:

$$f(\eta) = \frac{\int_0^{\eta} \frac{\mathrm{d}N}{\mathrm{d}\eta'}\,\mathrm{d}\eta'}{\int_0^{1} \frac{\mathrm{d}N}{\mathrm{d}\eta'}\,\mathrm{d}\eta'}\,, \tag{7.4}$$

where $\mathrm{d}N/\mathrm{d}\eta$ is the eta distribution [87].

The residuals give an indication of the performance of the sensor and the cluster position reconstruction algorithm, with a thinner residual distribution being more desirable, more details will be given in the next section. The residuals at perpendicular incidence for the two eta-correction implementation methods are shown in figure 7.7, along with the combination of the two methods. Comparing the standard deviation of the Gaussian fit, the second method is slightly better and further improvement can be made by combining the two, as described above.



(a) **(b)** **(c)**

Figure 7.7: The global x residuals at perpendicular incidence of the different eta-correction methods: (a) first, (b) second and (c) both of them together. A Gaussian fit is shown in red with the value of the standard deviation given.

7.2.2 Track reconstruction

The tracking works by defining a reference telescope plane, in this case the plane furthest upstream of the beam, from which the tracks originate. The track reconstruction algorithm individually looks at all clusters in the reference plane and for each one extrapolates to the next plane in the z- (beam) direction. It then looks for the nearest cluster on the next plane, and if this is within 200 µm then it is added to the track. Now the track contains 2 clusters. The track is then fit with a straight line using a linear regression method to get the best fit. This time, the extrapolation is along the straight line fit in the z-direction. This is done to the third plane and finds the nearest cluster. The algorithm repeats this until all telescope planes have been checked and fits a final straight line. For a track to be formed it must contain one cluster from each plane. The linear fit is utilised to determine the position on the DUT for the tracks. Only tracks with a value of χ^2/NDF < 3.5 are considered (see figure 7.8(a)), to remove tracks with significant scattering.

Once all tracks have been reconstructed their intercept with the DUT is calculated. To match a track to a cluster on the DUT, several checks are made. The first check is an association cut. The absolute difference between the global cluster position and the global track intercept, called the residual (see figure 7.8(b)), is required to be less than 100 µm. For all clusters, the match that has the smallest absolute residual value is used. Tracks are rejected if they went outside the fiducial region on the DUT, are adjacent to a masked pixel (± 1 in the x- and/or y-direction) or if the track position on the DUT is within a distance of 125 µm of another track.

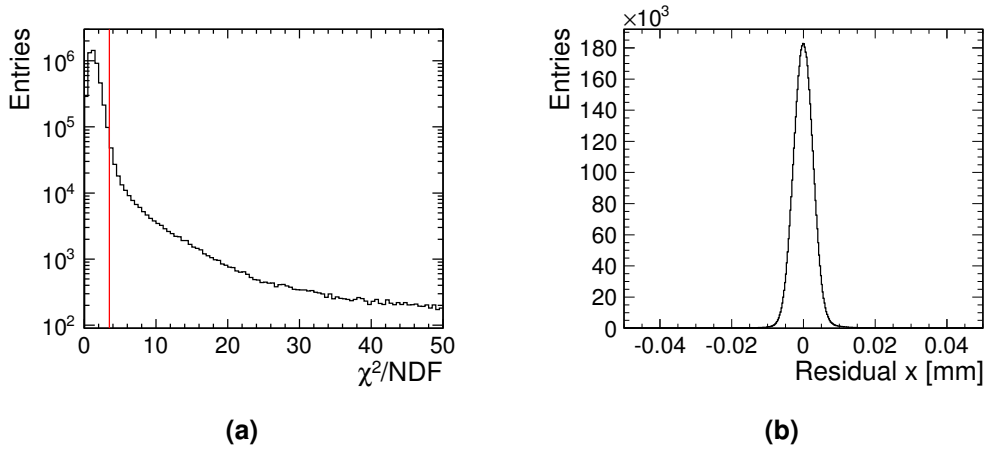


Figure 7.8: (a) The χ^2/NDF for all tracks on the device under test, showing the value of the cut at 3.5 and (b) an example of the residuals from one of the telescope planes.

7.2.3 Alignment

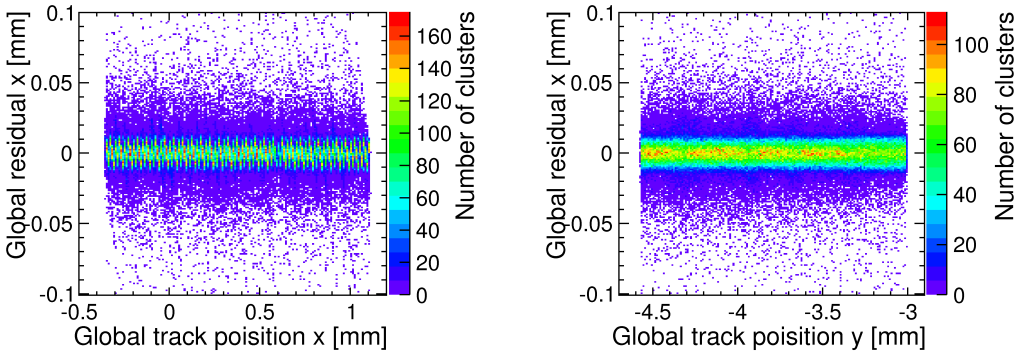
The sensors were placed so that their x-y plane was as close as possible to perpendicular to the beam and they had as little rotation as possible. In order to take into account small misalignments, a software alignment was applied. This uses the measured z-positions of the planes as a fixed parameter and adjusts the x and y coordinates as well as the rotations about each axis: θ_x, θ_y and θ_z, as these are the free parameters in the alignment. This was done offline and performed in two stages: first the telescope planes were aligned, then the DUT was aligned. A minimum number of 400k tracks was required to ensure an accurate alignment. The alignment of the telescope planes loops over all the planes, excluding the reference plane (the first plane in the telescope), and for each one, minimises the track χ^2 by varying the free parameters and fitting new tracks. It does this for all the tracks and once all the planes are aligned it will repeat the procedure 4 more times to ensure the calculated χ^2 is the minimum.

Once the telescope planes have been aligned the DUT is aligned separately with all telescope planes fixed. To align the DUT, the intersection point of each track with the DUT is determined by interpolation. In this case the residuals are minimised, varying the free parameters. By separating the alignment in such a way, it ensures an independent measurement of the track position on the DUT is achieved. To get a handle on the quality of the alignment, plots of the residuals as a function of the track position (track intercept on DUT) are used, examples are shown in figure 7.9. If the alignment is not done correctly then these plots will show a slope or an offset from zero. The abrupt edge to these plots occurs because tracks cannot travel outside the DUT. Another method is to check if the residuals (e.g. figure 7.8(b)) are centred around 0 in both the x- and y-direction.

7.3 Simulation

In order to compare the simulation results to data, a conversion from the collected signal to an observable ToT measurement on the CLICpix was required. This was performed in two stages. Firstly, the current pulse from TCAD, an example is given in figure 7.10(a), was fed into a circuit simulation of the CCPDv3 electronics[1]. This simulation was carried out using Cadence Virtuoso software [90], and allowed the analogue output of the CCPDv3 to be simulated, an example is shown in figure 7.10(b). The second stage was to use a calibration curve and a surrogate fit to the curve to convert this pulse height into a ToT. Details of how this curve and fit were obtained were given in chapter 6.3.

[1] This simulation was performed by I. Kremastiotis, CERN.



(a) (b)

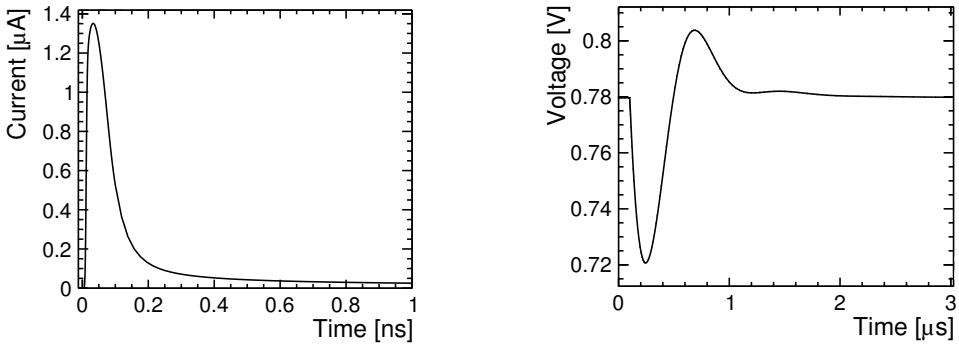
Figure 7.9: Examples of the plots used to assess the quality of the alignment. The x residual in the global coordinate system as a function of (a) the track x position and (b) the track y position.

7.4 Tracking performance

Once all the reconstruction algorithms are in place and all the data has been collected, the results of the beam tests can be produced. Details of the assembly's performance are collected such as the charge collection properties, the resolution and the efficiency. The results of the beam test can also be used to validate the simulations by comparing the two.

7.4.1 Charge collection

In order to compare the simulation results with data, artifacts of the detector assembly have first to be removed. The uniformity of the signal across the CLICpix depends



(a) (b)

Figure 7.10: (a) An example current pulse obtained from TCAD from a MIP simulation. (b) An example of the simulated analogue output of the CCPDv3 using a TCAD current pulse as input.

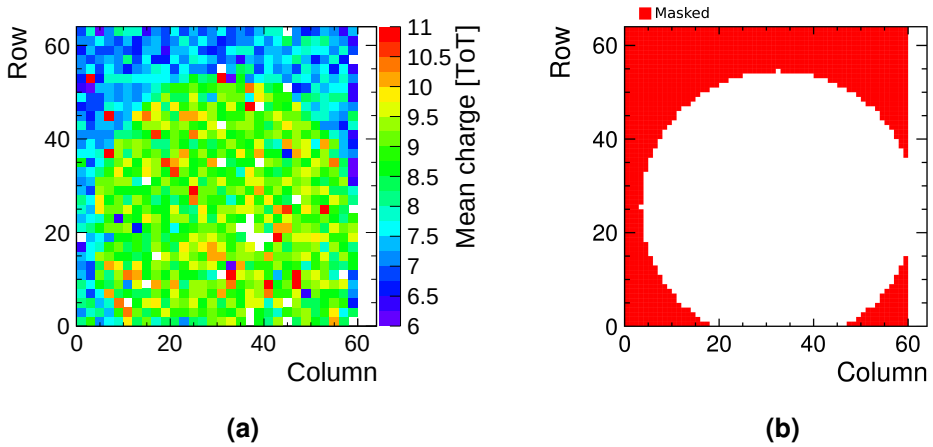


Figure 7.11: (a) Mean ToT per pixel averaged over 2x2 pixels at 0°. For each pixel position the ToT spectrum is fitted with a Gaussian to obtain the mean. (b) The circle mask applied with red indicating the masked pixels.

on the coupling to the HV-CMOS sensor; the response across the matrix is shown in figure 7.11(a). For each pixel, the ToT spectrum is fitted with a Gaussian distribution from which the mean is extracted and plotted in the figure. As can be seen, the matrix contains a clearly visible circle inside of which the pixels have a higher observed signal compared to the rest of the matrix. This is attributed to the deposited glue, which has a significant dielectric constant compared to air, not spreading fully across the chips. The region of lower signal has been neglected from the analysis and the subsequent mask applied is shown in figure 7.11(b).

In addition to the uniformity of the pixel matrix response, the relative alignment of the HV-CMOS and readout ASIC is an important factor in the measured performance. The coupling pads through which the signal is transferred from the CCPDv3 to the CLICpix are shown in figure 7.1, along with their alignment orientation for bonding. As the signal is transferred capacitively and the pads are large compared to the inter-pixel distance, capacitive coupling from a single HV-CMOS pixel to multiple pixels on the readout chip may take place. This has already been observed for this type of device [52]. An illustration of a cross coupled hit is shown in figure 7.12. The effect of cross-coupling is the same as cross-talk. However, since the CCPDv3 has a guard ring surrounding each pixel, the capacitance to the neighbours is low and cross-talk is not expected to be a major issue.

Hits due to cross-coupling can be observed by scanning the track intercept across an individual pixel and looking at the number of times that it causes hits in this pixel and its subsequent neighbours. This is seen in figure 7.13 where the difference between the track position on the DUT and the pixel centre of a pixel in the cluster, called track

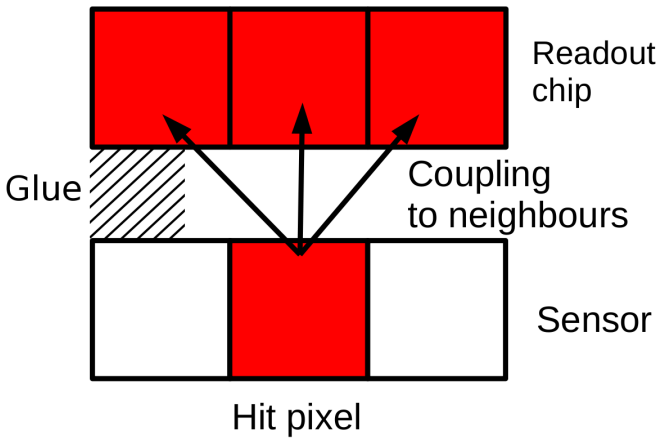


Figure 7.12: A sketch of a cross-coupled hit on the assembly.

distance, is plotted against the pixel response. The central peak is due to charge deposited inside the active volume of the pixel, while additional peaks outside of this range are assumed to be induced by cross-capacitances to neighbouring pixels which fire. These peaks are fit with a so called top hat function [52], which consists of the combination of two error functions. A central top hat fit for the hit pixel and additional top hat fits for subsequent neighbours is performed. The fit is given by:

$$\sum_{pixels} \frac{a}{2}\left[\text{erf}\left(\frac{x+x_{pixel}+\frac{w}{2}}{\sigma\sqrt{2}}\right)-\text{erf}\left(\frac{x+x_{pixel}-\frac{w}{2}}{\sigma\sqrt{2}}\right)\right], \tag{7.5}$$

where x is the track distance, x_{pixel} is the distance to the neighbouring pixel border, w is the width of the plateau of the top hat and σ is the width of the error function [52]. Both the width, w, and the sigma of the error function were left floating while the offsets to the neighbouring pixel borders, x_{pixel}, were fixed. The outcome of this fit is shown in red in figure 7.13 and the top hat functions for the hit pixel and neighbours either side are shown in blue. The observed asymmetry seen in figure 7.13(b) is attributed to the CLICpix pad not being symmetric this direction, as seen in figure 7.1.

In order to include the effects of cross-capacitance in the simulated detector response, the relative magnitude of the pixel-to-neighbour capacitance, with respect to the central coupling, was estimated. To obtain this number, several steps needed to be made. Firstly, an estimate of the fraction of hits which arise due to cross coupling was performed. This is done by only looking at clusters with cluster x-width 1 and cluster y-width 2 (1x2) for odd columns, as this would simplify the model and negate the charge injection issue by only looking at one type of column. The residuals of such clusters were then fit with two Gaussians, a broad one for the suspected cross-coupled hits and

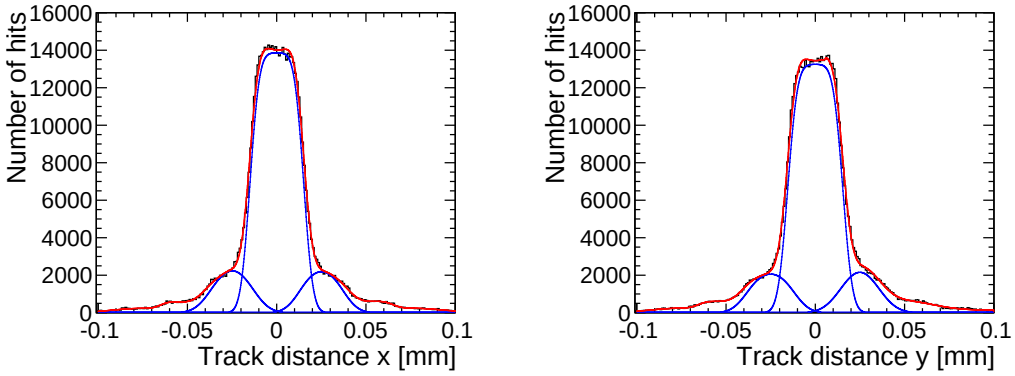


Figure 7.13: Pixel response plot as a function of the distance of the track position from the pixel centre for tracks at perpendicular incidence for the (a) x- and (b) y-directions. The total top hat fit is in red while individual top hat fits for three pixels are in blue.

a thin one for the diffusion charge sharing hits. Figure 7.14(a) shows these two fits in blue and the combination of these two fits in red. From these fits, the amount of cross-coupled hits is given by a ratio of the integral of the broad distribution to the integral of the total fit, yielding a value of $\approx$ 33%. Then the cluster y-width distribution for odd clusters, figure 7.14(b), is used to get the number of clusters with a cluster y-width of 2. Multiplying this by $\approx$ 33% produces the number of cross-coupled clusters for clusters of cluster y-width 2. Dividing this by the total number of clusters with cluster y-width 1 or 2 gives the number of cluster y-width 1 clusters that caused cross-coupled hits, i.e. an estimate on the number of hits that arise from cross-coupled hits. This value is $\approx$ 20%.

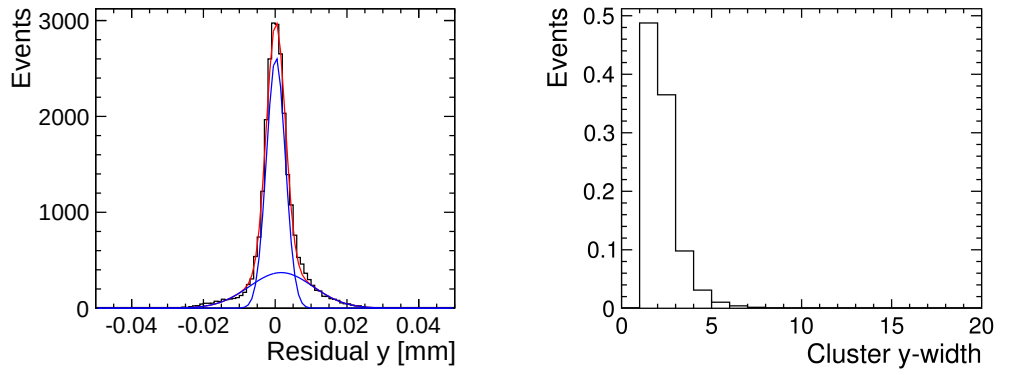


Figure 7.14: (a) The residuals of clusters with cluster x-width 1 and cluster y-width 2 for odd columns. The red fit is the sum of the two Gaussian fits shown in blue. (b) The cluster y-width distribution for odd column clusters.

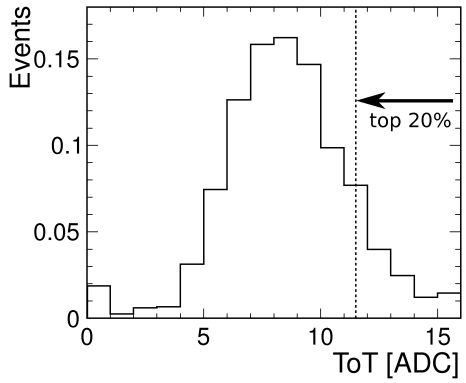


Figure 7.15: The hit pixel ToT spectrum for tracks passing through a central square of width 3 µm. The vertical dashed line represents the estimated threshold for cross-coupled hits.

The next step is to assume that the cross-coupled hits are caused by the hits with highest magnitude in the hit pixel charge spectrum for tracks passing through a central square of width 3 µm [1]. By integrating the histogram of the hit pixel charge spectrum from the upper energy limit until 20% of pixels has been covered, the energy threshold for producing a cross-coupled hit to the non-diagonal neighbours can be estimated, producing a ToT value of 11. The ratio of this threshold to the data-taking threshold of the matrix will then be the ratio of the cross-capacitance with respect to the central capacitance. From the hit pixel charge spectrum shown in 7.15, this results in a value of $\approx 4.5\%$. From simulations studies carried out, the largest capacitance to a neighbour pixel, relative to the main coupling, was 3.7% [91], giving good agreement between the two.

To implement the estimated cross-coupling value into the simulation the simulated pulse heights were manipulated. This was done by taking the value of the pulse height of the adjacent pixel at the same in-pixel position, i.e. either ± 25 µm, then multiplying this by 5% and adding it on. The result of this can be seen in figure 7.16. It shows the simulated pulse height as a function of track position with no cross-coupling, a cross-coupling of 5% added and only the 5% cross-coupling contribution. The effect of the cross-coupling is to increase the pulse height and raise it above the horizontal dashed line which represents the threshold obtained from the calibration curve. The cross-coupling increases the response width of the simulation by around 0.01 mm either side of the pixel. It has a minimal effect on the pulse height within the pixel cell as the pulse height outside is very small.

[1] This central region was chosen as it would be more likely that a multi-pixel hit would be from cross-coupling than diffusion.

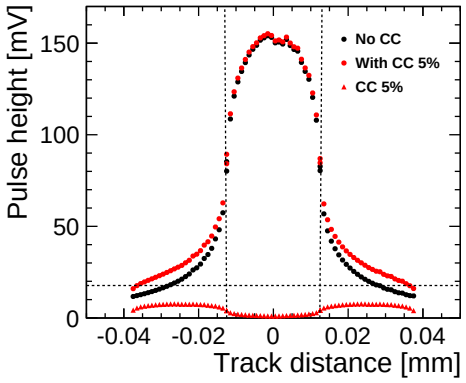


Figure 7.16: The simulated pulse height of the CCPDv3 as a function of track distance. The vertical dashed lines represent a one-pixel cell and the horizontal dashed line represents the threshold.

The cumulative effect of these additions can be seen in figure 7.17, where a comparison of the pixel ToT as a function of track distance is shown for both simulation and data. Data is shown for clusters in the x-direction of the CLICpix and for perpendicular track incidence, where the track passes close to the centre of the pixel in order to be represented by the 2D TCAD simulation. There is a good agreement between simulation and data, despite several factors which are not taken into account in the simulations, particularly Landau fluctuations of the deposited charge.

7.4.2 Bias scan

An advantage of using TCAD descriptions of the sensor electric field for simulation, is the ability to describe the detector response over a wide range of bias voltages and

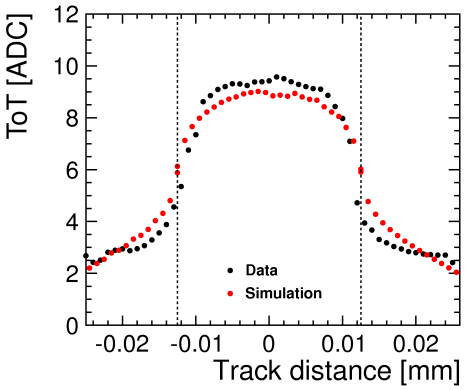


Figure 7.17: Pixel ToT as a function of the distance of the track position from the pixel centre in the column direction, for data (black) and simulation (red). The vertical dashed lines represent a one-pixel cell.

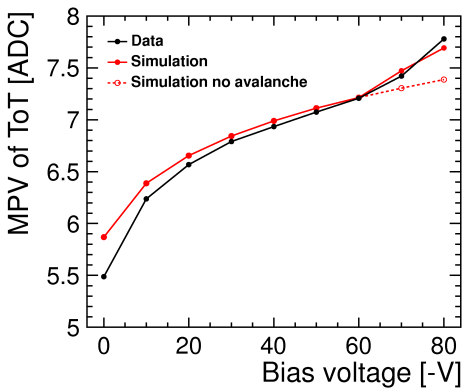


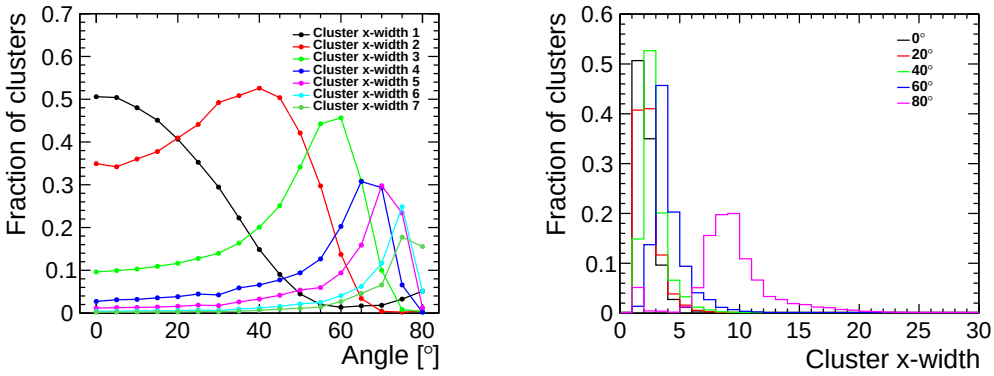
Figure 7.18: Most probable ToT value for single pixel clusters as a function of bias at perpendicular incidence, for data (black) and simulation (red). The dashed red line represents the simulations without avalanche model.

field configurations. A comparison of results for different bias voltages is shown in figure 7.18, where the most probable value of the ToT distribution for single pixel clusters at perpendicular incidence can be seen for simulation and data. While both exhibit the typical increase in collected charge expected from the extension of the depletion region between 0 and 60 V, the data beyond this range shows a marked rise in the amount of charge measured by the sensor. This rise is only reproduced in simulation with the addition of an avalanche model to TCAD, which reproduces the observed behaviour. The avalanche model used was the University of Bologna impact ionization model for silicon [92]. High field regions in the vicinity of the deep n-well thus produce avalanche conditions for the multiplication of charge close to the implants. Between the data and the simulations there is an agreement of better than 10%. Due to the sensor breakdown at –93 V more data points could not be taken in order to make a stronger case.

7.4.3 Charge collection for angled tracks

Results presented so far have been for tracks at perpendicular incidence to the detector surface. For the proposed CLIC vertex detector however, track angles of up to 75° are expected [28]. The response of capacitively coupled assemblies are studied over a range of angles to better understand their performance. The rotation was done around the y-axis, figure 7.2. The beam test results have not been optimised as they do not take into account the effects that occur when going to large angles, such as the impact on track-cluster association or split clusters.

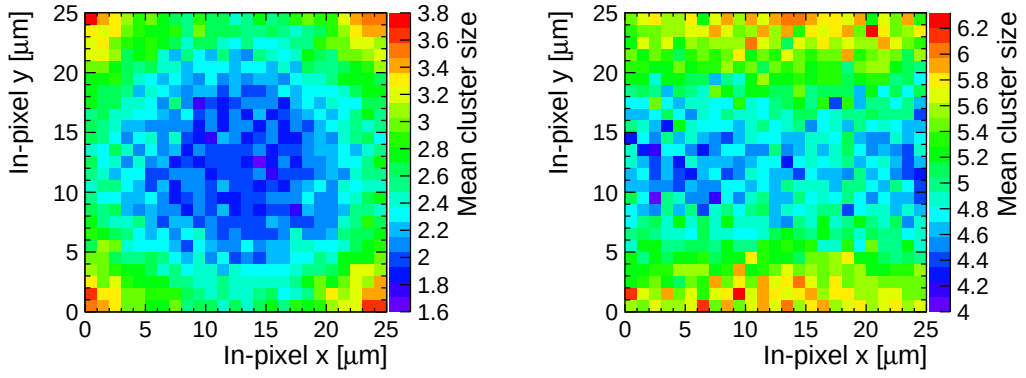
The cluster x-width distributions for an array of angles is shown in figure 7.19, along with the fraction of pixels with cluster x-width n as a function of angle. Given the limited depth of the depleted region (simulated value of 11.5 µm at –60 V), even at angles of



(a) **(b)**

Figure 7.19: (a) Fraction of pixels with cluster x-width n as a function of angle. (b) cluster x-width distributions in the direction of rotation, for various angles.

up to 60°, the majority of clusters display widths of less than 5, while those produced by tracks almost parallel to the detector surface have a large variance in their observed size. This variance can be attributed to several effects: δ-rays, edge effects and split clusters caused by masked pixels. Looking in more detail at the cluster size over the pixel cell, the charge sharing behaviour of the sensor can be observed. Figure 7.20 shows the mean cluster size as a function of in-pixel track position for data at 0° and 60°. At perpendicular incidence, the mean cluster size in the centre of the pixel takes values of around 2. This can be attributed to the tail of the charge distribution and the small pixel size generating many multi-pixel clusters. For a track angle of 60°, the cluster size is independent of the x-direction because of the shallow incident angle.



(a) **(b)**

Figure 7.20: Mean cluster size as a function of in-pixel track position for (a) 0° and (b) 60°.

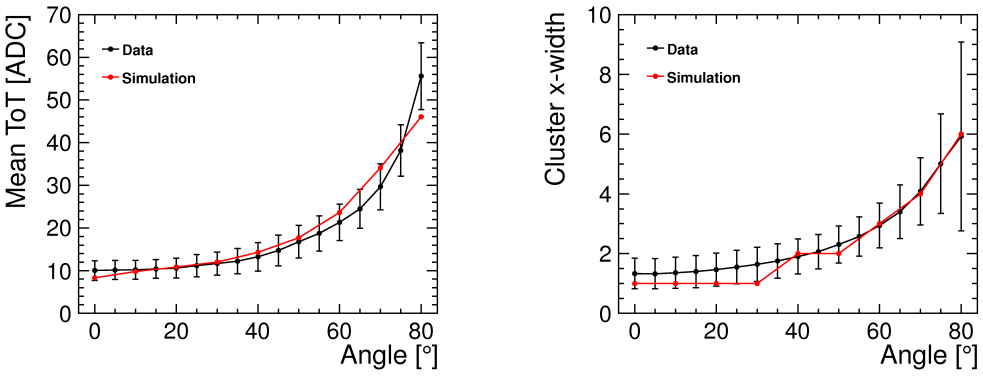


Figure 7.21: (a) Mean value of the cluster ToT distribution as a function of angle. (b) Mean cluster x-width as a function of angle. The vertical bars denote the RMS of the histograms. Simulated values are shown in red.

The mean value of the cluster ToT (sum of the individual pixel ToTs in the cluster) and the mean cluster width over the angular range are shown in figure 7.21, with the vertical bars denoting the RMS of the histograms. The simulations are in good agreement with the beam test data even at the higher angles. In both cases, the results are restricted to clusters with a cluster y-width of one and an additional cut on the track position within $\pm 2.5\,\mu m$ of the pixel centre was made. For the mean charge, this was applied in both x- and y-directions, while for the mean cluster x-width this cut was only applied to the y-direction, creating a strip in the x-direction. This was done to better replicate the 2D nature of the simulations, allowing for a better comparison between the two.

7.4.4 Active Depth

The timing characteristics of charge collected from the depleted and undepleted regions of the bulk were shown in chapter 5.6. A measurement of both of these contributions comes through an estimation of the active depth. This is the depth at which charge is collected and contributes to the signal. The active depth of the sensor will be sensitive to additional factors such as the integration time of the electronics and the operating threshold of the readout ASIC. In order to get an estimate of the active depth, a geometric approximation is taken which makes use of the known parameters including the pitch, p, and the rotation angle, θ. These quantities can be related to the cluster width in the direction of rotation by the following equation:

$$\text{cluster x-width} = \tan(\theta)\frac{d}{p} + c, \tag{7.6}$$

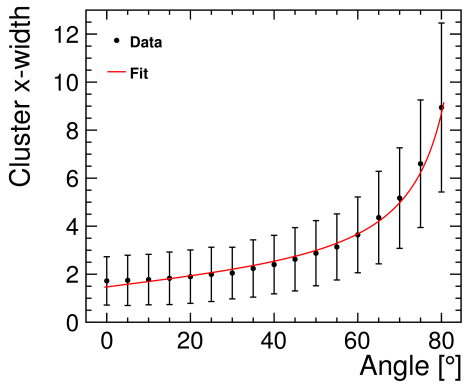


Figure 7.22: The mean cluster x-width as a function of rotation angle with the fit (equation 7.6) shown in red. The vertical bars denote the RMS of the cluster x-width distributions.

where c is the cluster x-width at perpendicular incidence, and d is the active depth. This is used to fit the data in figure 7.22, which shows the mean cluster x-width as a function of angle with the RMS denoted by the vertical bars. From this, a value of $\approx 30\,\mu m$ was found for the active depth - almost 3 times as large as the simulated depletion depth of 11.5 µm given in chapter 5.5.2 - indicating a significant contribution from diffusion to the charge collection, in agreement with results in chapter 5.6. The result is also in agreement with the measured active depth of 25 µm [78].

7.4.5 Single hit resolution

One of the challenging requirements for the CLIC vertex detector is to provide a single hit resolution of 3 µm, while retaining time-tagging precision of $< 10\,ns$, chapter 4.1. The single hit resolution of the capacitively coupled CCPDv3-CLICpix assembly has been measured over the full angular range. By looking at the residuals the resolution can be obtained. The measured residuals are related to the telescope resolution and the single hit resolution on the DUT in the following way:

$$\sigma^2_{\text{residual}} = \sigma^2_{\text{DUT}} + \sigma^2_{\text{telescope}}\,. \tag{7.7}$$

The resolution of the telescope depends on many variables such as: single point resolution of the sensors in the telescope plane, the number of measurements, the position of the planes and the amount of multiple scattering [93]. In this study the resolution of the DUT is defined as the standard deviation of a Gaussian fitted to the residual distribution for all clusters. Therefore, the telescope resolution has not been subtracted in quadrature from the measured value but will not significantly impact the values because it is small ($\approx$1.6 µm).

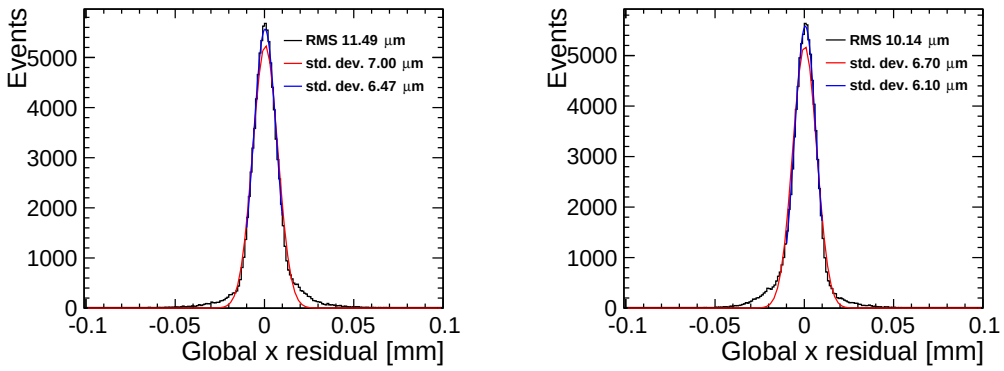


Figure 7.23: Residuals in the x-direction before the eta correction at (a) 0° and (b) 80°. The red line represents a Gaussian fit between ± 0.1 mm, while the blue line is for a Gaussian fit between ± 10 µm.

Figure 7.23 shows the residual distribution for all clusters at 0° and 80°, before the eta corrections. A Gaussian fit, shown in red, within ± 0.1 mm, has been made and is a good approximation of the residual. Comparing the standard deviation of this fit to the RMS of the histogram, it can be seen that the fit produces a better result for the resolution. Reducing the range of the Gaussian fit to ± 10 µm produces a further improvement and allows to better fit the peak. This indicates that some negative effects such as cross-coupling have an impact on the residuals. For the resolution values used in this study, the red fit range was used.

Breaking down the residuals into the individual residual distributions for perpendicular track incidence, both before and after the application of eta corrections to account for the non-linear sharing of charge between pixels, is shown in figure 7.24. While the residuals improve sharply as expected, the effects of multi-hit clusters due to capacitive cross-coupling can be clearly seen. When considering clusters with cluster x-width 2, two distinct distributions can be seen: those which are the result of tracks close to the pixel boundaries and where the charge in each pixel is proportional to the track position, and those where the second pixel has fired due to cross-coupling. In the latter case, the position resolution of the cluster can only be degraded, and a much wider residual distribution, even than for single pixel clusters, is observed. The asymmetry seen in the multi-pixel residuals could be due to asymmetry in the CLICpix via in both the x and y direction (figure 7.1). It could also indicate that there is a small misalignment between the coupling pads. Since an exact measurement of this alignment is only possible through destructive processes, the alignment obtained during the coupling is a best estimate, which has an error of ≈ 2 µm. Another source of the wide residuals in

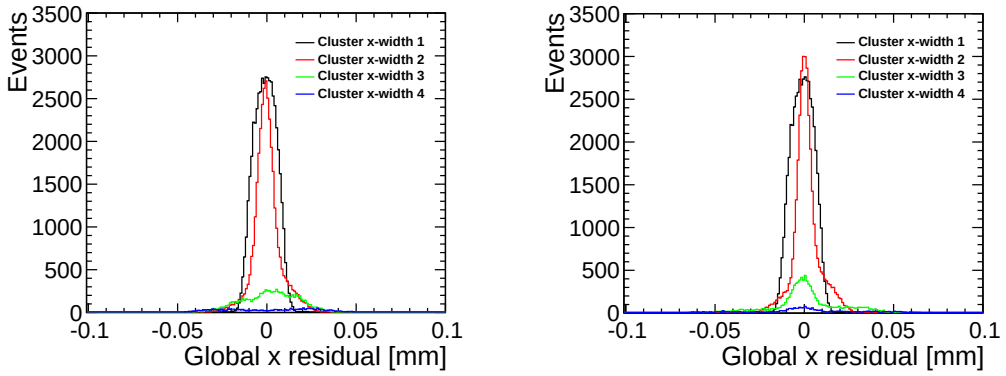


Figure 7.24: Residuals in the x-direction at perpendicular incidence for different cluster x-widths (a) before and (b) after eta correction.

the multi-pixel hits, are hits due to delta rays which lead to cluster depositions that do not correlate well with the incident position of the track. It is interesting to note that this asymmetry in the residuals is not present when the circle mask is not used but these could be hidden by the worse response from pixels outside the mask.

The measured resolution can be seen versus angle, both before and after eta corrections, in figure 7.25. The reason for the eta correction making the resolution worse than the value for the centre of gravity algorithm at 80°, is that the correction is done in local coordinates where the track's local coordinates are not well defined and increase the potential for a mismatch between the centre of gravity position and the track position. At 60°, the eta correction produces the largest reduction in the resolution, after this it becomes worse. To avoid this, a new cluster position algorithm can be used, one that is more suited to large angles, the head-to-tail algorithm [88]. However, this was not investigated in this study.

7.4.6 Single hit efficiency

Of importance for all particle physics experiments is the detector efficiency, which should generally be greater than around 99%. The efficiency for capacitively coupled assemblies has already been measured previously [52], but is presented here over the full angular range. The fiducial cuts described in chapter 7.4.1 are applied in figure 7.26 which shows that the device tested had efficiencies of $>$99.8% over the full angular acceptance, with a small improvement at higher angles where several pixels are crossed. The value of the efficiency at 30° is not consistent with the other angles because for some runs the beam spot was not aligned resulting in low number of events for several runs. Figure 7.27 shows that across the chip and within the pixel cell, split into odd and

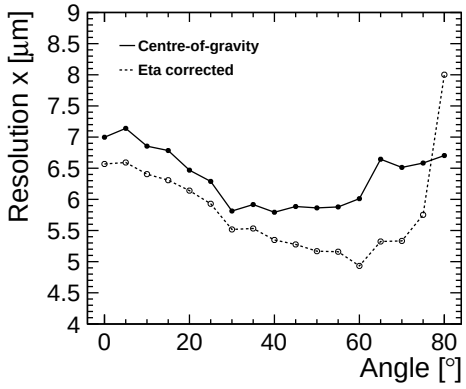


Figure 7.25: Single hit resolution of all clusters combined in the x-direction at different angles, for centre-of-gravity (bold) and eta-corrected (dashed) cluster positions.

even columns, the efficiency stays mainly uniform. Across the matrix the areas with slightly lower efficiency are those that are outside the fiducial region and have slightly lower coupling due to the absence of the glue, as discussed in chapter 7.4.1. The total efficiency outside the fiducial region is 99.5 ± 0.03, remaining above 99%.

7.4.7 Effects of cross-coupling

The major effect of cross-coupling is to spoil the results of the position finding algorithm and hence the resolution by artificially increasing the cluster size and cluster charge. As mentioned in chapter 7.4.5, hints of the cross-coupled hits are seen in the residuals, cluster with values greater than $\pm\,0.01$ mm. To substantiate this statement further, a dedicated analysis on the effects of cross-coupling has been performed. This was done by studying clusters with a cluster x-width of one and a cluster y-width of two

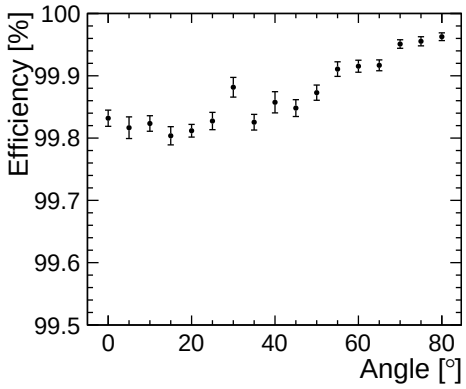


Figure 7.26: Single hit efficiency of the DUT as a function of the track angle of incidence with the error bars denoting the statistical error on the mean value.

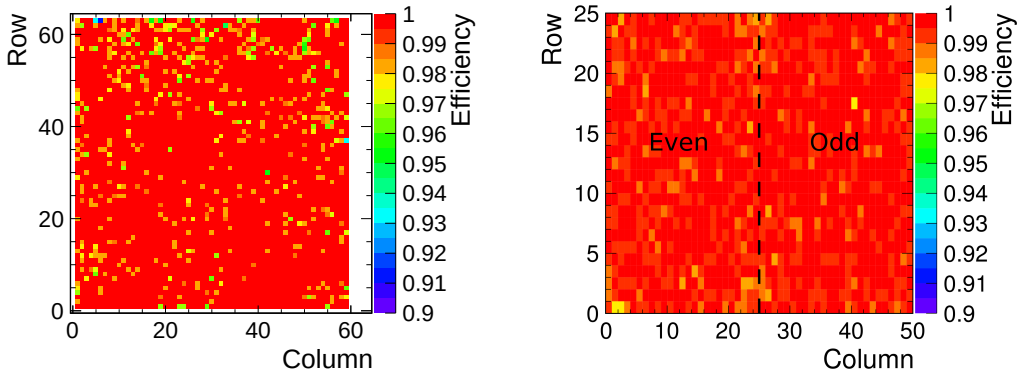


Figure 7.27: (a) The efficiency across the matrix and (b) the efficiency within the pixel cell split into odd and even columns.

(1x2 clusters), to potentially isolate cross-coupled hits as much as possible, or at least increase the likelihood of a cross-coupled hit occurring. As described in chapter 7.4.1 the cross-talk is expected to be very small so coupling to neighbours is very likely due to capacitive cross-coupling. To avoid the charge injection issue in the CLICpix described in chapter 6.2.3, clusters with cluster x-width two and cluster y-width one were not considered and only one type of column, odd, were included. The following results are for cluster positions given by the centre of gravity algorithm.

The first investigation was to examine the effect of cross-coupled hits on the residuals in more detail by looking at how the residuals depend on the track's in-pixel y-position, as shown in figure 7.28. For hits that are at the pixel boundary, either at an in-pixel y position of 0 µm or 25 µm, the majority of the hits for the 1x2 clusters should come from charge sharing. As a result, the residuals are thin and centred around 0 because the algorithm can correctly reconstruct the track position. An illustration of this is shown in figure 7.29(a) and the results of a projection along the x-axis of figure 7.28 at a track in-pixel y position of 0 µm are shown in figure 7.30(a). Even at this track position, a small amount of cross-coupling is seen at ± 0.025 mm.

For the second case, when the track passes through the centre of the hit pixel, an assembly without cross-coupling would not produce many 1x2 clusters. While for an assembly with cross-coupling it can be estimated that a considerable number of these hits are due to cross-coupling. This causes the cluster position to be closer to the boundary between two pixels and further away from the track position, as shown in figure 7.29(b). This produces bad residuals, figure 7.30(b), with two peaks at ± 0.01 mm. This is consistent with what was seen in figure 7.24(b) for a cluster x-width of 2, where there are two distributions for the residual, a thin distribution and a broader distribution,

this was explained in chapter 7.4.5 .The appearance of two distinct distributions confirms that hits with residuals outside of ± 0.01 mm are likely due to cross-coupled hits. The asymmetry seen in figure 7.30(b) possibly arises from the CLICpix pads not being symmetric in y, as mentioned in chapter 7.4.1. This study is still not able to distinguish the cross-coupled hits from poorly reconstructed hits that result in residual values of ± 0.01 mm. Further evidence is needed.

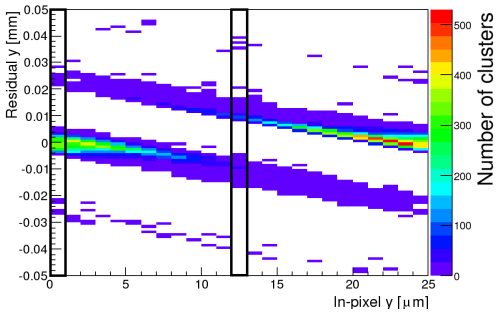


Figure 7.28: The residuals in the y-direction as a function of track in-pixel y-position for 1x2 clusters. The two projections along the x-axis are indicated by the black boxes.

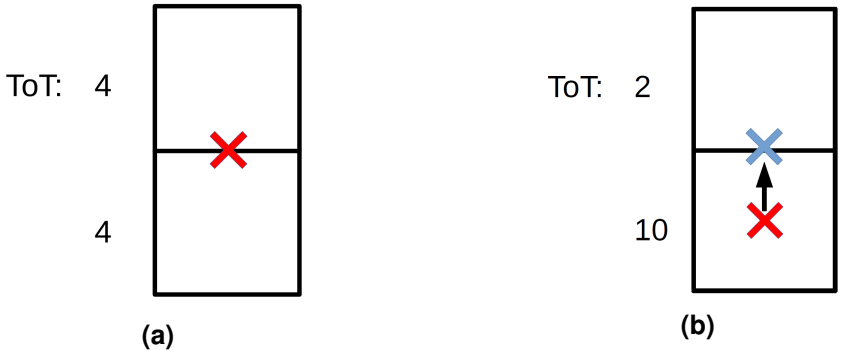


Figure 7.29: Illustration of the effects of cross-coupling on the GoG position (blue) for two different track positions (red): (a) boundary between two pixels and (b) the centre of the pixel.

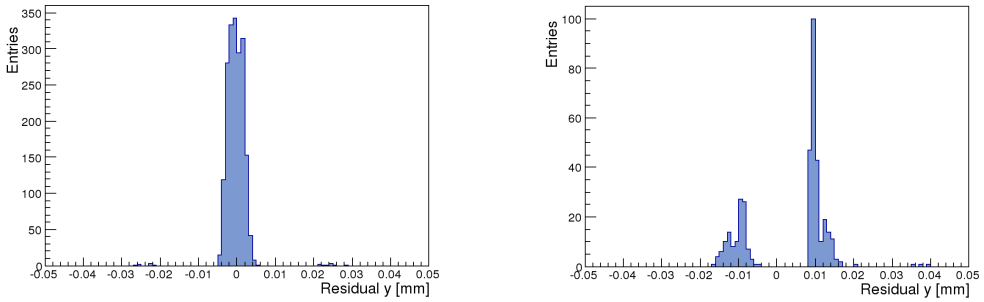


Figure 7.30: Residuals at different track in-pixel y positions: (a) 0 µm and (b) 12.5 µm.

Another effect that the cross-coupled hits can have, is to increase the ToT of the cluster. Using the knowledge of the residuals of cross-coupled hits detailed above, a cut can be made on the residuals. This cut was made to separate clusters with residual values within ± 0.01 mm, from those outside this range, as shown in figure 7.31. Using these two types of clusters, the cluster ToT was compared, figure 7.32. There is a distinct difference between the two, with the clusters that lie outside this range showing two peaks in their cluster ToT. This is due to the addition of the induced charge onto the neighbouring pixel and the fact that cross-coupled hits occurs when the pixel has a high ToT, greater than 11 as described in chapter 7.4.1. Another possible reason for this is δ-rays, but these are rare events. An estimate of the number of cross-coupled hits can be made by finding the integral of the cross-coupled bump, in the range 21-33, and dividing this by the total integral, producing a value of $\approx 36\%$. This is slightly higher than the $\approx 33\%$ given in chapter 7.4.1 but is to be expected as the tail of the distribution in figure 7.32(a) bleeds into the cross-coupled bump.

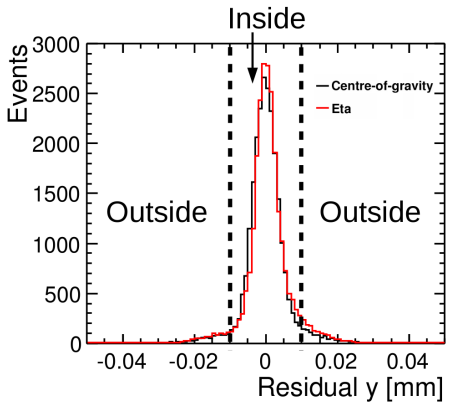
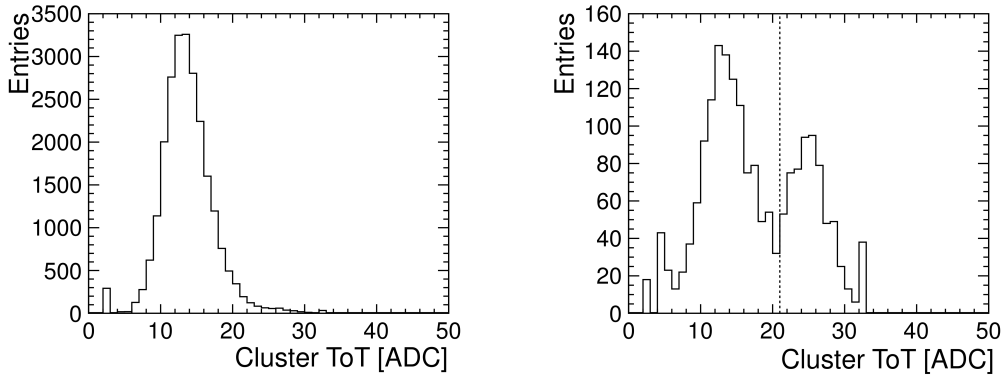


Figure 7.31: Residuals of 1x2 clusters illustrating a cut made on the residuals to distinguish the cross-coupled hits.



(a) **(b)**

Figure 7.32: Cluster ToT of 1x2 clusters (a) inside and (b) outside the residual cut. The dashed line represents the boundary of the cross-coupled hits.

To further demonstrate the link between a high pixel ToT and a cross-coupled hit, the largest pixel ToT of a 1x2 cluster with a track position within a central 7x7 square was plotted as a function of y residual. These types of clusters are expected to be produced by cross-coupled hits. For this reason their residuals should be at ± 0.01 mm as seen in figure 7.33. In addition to this, the largest pixel ToT should be above the cross-coupled threshold of 11 estimated in chapter 7.4.1, but this is not true for all clusters shown. This could indicate that there are still some non cross-coupled hits in the figure, even after the extensive cuts used, that have poor residual values. It could also indicate that the estimate of the cross-coupling threshold is too high. However, the mean largest ToT value is roughly the same, within 3%, as the estimated cross-coupling threshold so a large fraction of the clusters shown can be attributed to cross-coupled hits. The difference could be due to the fact that the cross-coupling threshold was calculated using clusters within a central 3x3 square instead of central 7x7 square, so the mean would increase if central 3x3 square had been used in figure 7.33.

From all the above studies into the effects of cross-coupled hits, a general trend is shown: if a track goes through the centre of the pixel and the resulting cluster has both a residual larger than ± 0.01 mm and the largest ToT is greater than 10, then it can be said that this cluster will suffer from cross-coupling. However, it cannot be said for certain that this type of cluster is a cross-coupled hit, only that it is more likely to be one. Hence, completely disentangling the cross-coupled hits will require more investigation. Nevertheless, it could be possible to use the ToT information alone, since the residual is not known in the real detector, in view of improving the reconstructed cluster position by using the pixel centre of the largest ToT pixel if the cluster meets the above criteria.

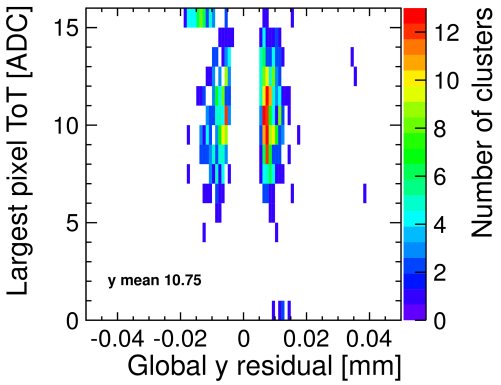


Figure 7.33: The largest pixel ToT of a 1x2 cluster with a track position within a central 7x7 square plotted as a function of y residual.

7.5 Summary

Measurements of HV-CMOS assemblies for the CLIC vertex detector have shown excellent tracking performance across the full angular acceptance. The single hit efficiency is above 99% and the single hit resolution is within 5–7 µm after eta correction. For the device under test used in this study the coupling between the HV-CMOS sensor and readout ASIC was not uniform across the matrix, owing to the limited amount of glue used during the assembly production. Despite this, a reasonable amount of coupling existed between the sensor and readout chip in places where the glue was not present, the efficiency is above 99%.

TCAD simulations combined with source calibrations have been used to reproduce the test-beam performance of the sensor, showing good agreement with measurements of charge collection characteristics. The main limiting factor on the performance of the capacitively coupled assembly is cross-coupling, impacting on the achievable resolution. To further improve the resolution, the cross-coupling must be reduced as much as possible. One method to do this, is to reduce the size of the metal coupling pads, through which the signal is transferred, with respect to the pixel pitch. Another method, that has been implemented on the next generation of chips, is to create a guard ring around the metal coupling pads on both the sensor and readout chip. Simulation studies have shown that this produces a reduction in unwanted cross-coupling by a factor of 10, without impacting the main coupling [91]. These guard rings have been implemented for the C3PD and the CLICpix2 which are currently under investigation [61].

CHAPTER 8
Conclusions

The physics goals set out by CLIC will require a vertex detector that is able to provide precise measurements, as well as be able to cope with the environmental conditions. As the vertex detector is closest to the interaction point, it has the most challenging demands imposed on it. To reach these goals a R&D program is being performed and this thesis has contributed to this by the development and evaluation of HV-CMOS pixel sensors and planar pixel sensors for the CLIC vertex detector.

One part of the R&D programme is simulations of the sensor, which have been carried out with the TCAD software, allowing detailed knowledge of the semiconductor properties of the CCPDv3 HV-CMOS sensors to be acquired. The first goal was to test the validity of the 2D simulation by comparing them to the more resource consuming 3D simulations. The results of this comparison showed that for the electric field, current-voltage, and charge collection after 100 ns (the simulated time at which the charge collected contributes to the signal), the full 2D and reduced 3D models were consistent within 15%. However, the capacitance-voltage showed a difference of 40%. This indicates that the 3D simulations are better suited to the capacitance measurements. Utilising that the 2D model, a 2D three pixel model was created and showed a good agreement, within 5%, between the simulated current-voltage characteristics and the measurements. It was also used to demonstrate that the total charge collection across the surface of the structure was approximately uniform.

As well as giving insights into the sensor properties the simulations can be used to gauge the possible prospects for improved performance. This was done in two ways: by using different bulk resistivities and biasing from the backside. By using higher resistivities the performance of the sensor was improved: an increase in 10 V to the breakdown voltage, an increase in 15 µm to the depletion depth at –60 V, a reduction in capacitance of 2 fF at –60 V and an increase in collected charge of 1000e^- after 100 ns. The only exception when using higher resistivities is that the leakage current increases by a factor of 2 at –60 V. These improvements are amplified when biasing from the backside, particularly for the 1 kΩ cm model, for example the depletion depth increased by 55 µm at –60 V. However, biasing from the backside comes with additional complexities and cost. Since the sensor may be operated below the depletion voltage, it is important to understand the relative contributions from the two types of charge collection: drift and diffusion. Splitting the MIP path allowed this to be done and found that diffusion becomes dominant after 50 ns for the 10 Ω cm model.

Besides capacitively coupled active HV-CMOS sensors, bump-bonded planar sensors are also considered for the CLIC vertex detector. Bump-bonding for a small pitch of 25 µm on the chip-to-chip level is currently not industry standard; new techniques need to be used. Therefore, bump quality measurements of the planar assemblies were carried out and various bump bond failure modes were identified. The results were used to identify pixels with acceptable bump bond quality, amounting to between 46% and 97% for the four investigated assemblies.

On top of this, calibrations were also carried out to convert the ToT charge measurements into physical units, such as collected electrons. The test pulse circuitry found in each pixel of the readout chip was used for this to good effect. Another type of calibration was used to convert the DAC voltage steps, used to control the chip threshold, into physical units. This was carried out with the test pulse injector and produced a value of (11.2-12.2) e^- per DAC step, reproducing the design value of 10 e^- within 20%. As with the test pulse ToT calibration, it relied on knowledge of the test pulse capacitance. As a result of these lab measurements, several issues were addressed with the design of the new readout chip, the CLICpix2.

For the HV-CMOS assemblies, ToT calibrations were attempted using a ^{90}Sr radioactive source. This took advantage of the analogue readout for some pixels in the sensor. It permitted a comparison between the output of the amplifier, found in each pixel on the sensor, with the ToT response of the CLICpix. This curve would later be used for the conversion of the simulated pulse height to a ToT for comparison with the beam test data.

Another method to test the prototype sensor assemblies is to conduct beam test measurements. This was done with the HV-CMOS assemblies over the full angular acceptance, recording an excellent tracking performance. A single hit efficiency of above 99% and a resolution of 5–7 µm after eta correction were found at all angles. This value is still not at the CLIC target of 3 µm due several factors. One such factor is the charge interpolation being limited by the fake multi-pixel hits that emerge due to cross-coupling. Another impact on the performance was the gluing process, where evidence of a glue spot not spreading was found in the assembly. This produced a non-uniform response in the matrix. Furthermore, the resolution is limited by: the small amount of charge sharing due to an active depth of only $\sim$30 µm, the threshold used was quite high and the charge injection issue of the CLICpix. The beam test data was compared to TCAD simulations where applicable and there is a good agreement between the two. The simulations were able to reproduce the charge collection as a function of angle.

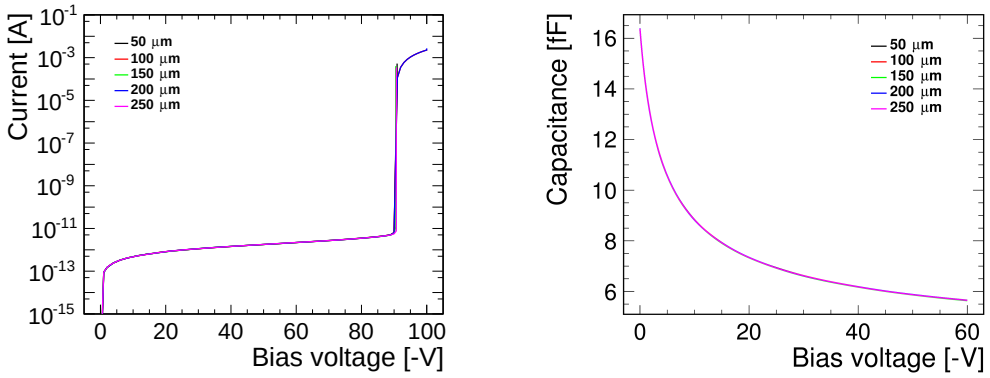
Altogether the studies in this thesis have produced a better understanding of how the proposed sensor technologies compare to the goals set out for the vertex detector. This knowledge can then be used to improve the next generation of chips and help to progress the vertex R&D for CLIC.

Appendices

APPENDIX A

Effect of simulation thickness

The choice of thickness for the studies in chapter 5.2 was 100 µm. For the real sensor the thickness is 250 µm and the target for CLIC is 50 µm. This simulation thickness was chosen to limit the mesh for the reduced 3D model when simulating a MIP. It was estimated that over the time scale that the real assembly would operate, a few hundred ns, the thickness would not impact the simulation results. To verify this simulations of the structure were carried out at five thickness: 50 µm, 100 µm, 150 µm, 200 µm and 250 µm. Figure A.1 shows a comparison between the current and capacitance for the various thicknesses, in both cases there is very little difference between the various cases. A difference begins to arise when looking at the charge collection over long time scales, as shown in figure A.2. Up until 600 ns there is no difference between 100 µm and 250 µm and even after 1 µs the difference is around 3%. Once 10 µs has been simulated a clear pattern begins to emerge, the larger thicknesses collect more charge. This is expected as there is a larger bulk due to the larger thickness, hence a larger contribution from the slower diffusion.



(a) (b)

Figure A.1: Current and capacitance as a function of voltage comparing different thicknesses.

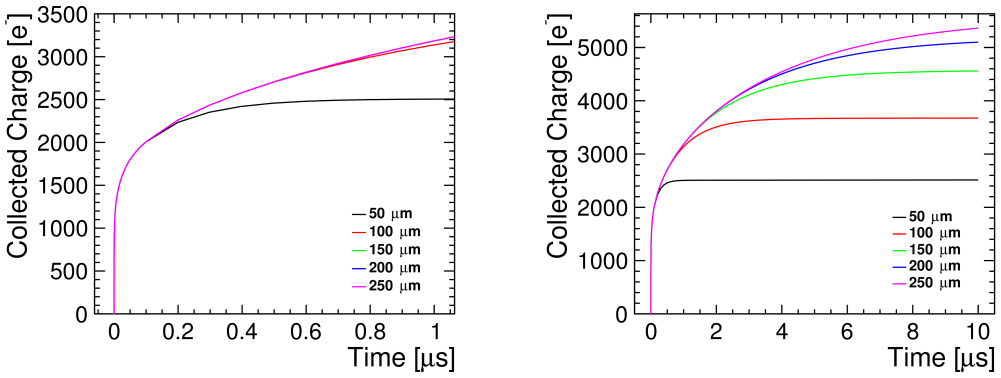


Figure A.2: Collected charge comparing different thicknesses and time scales: (a) 1 µs and (b) 10 µs

APPENDIX B

Reduction in capacitance

In chapter 5.5.4 the capacitance simulations for higher resistivities showed a sudden drop in capacitance. This was not anticipated but was guessed to be related to the depletion regions of neighbouring pixels joining together or reaching the edges of the simulation. To study this, the three-pixel model was slightly wider, 81.5 µm compared to 75 µm. This was to ensure a distinction between the depletion region of neighbouring pixels joining and the depletion region touching the edge of the simulation. As shown in figure B.2 this effect is due to the depletion region touching the edges of the simulation and, as shown in chapter 5.5.4, the PBC do not nullify this artefact of the simulation.

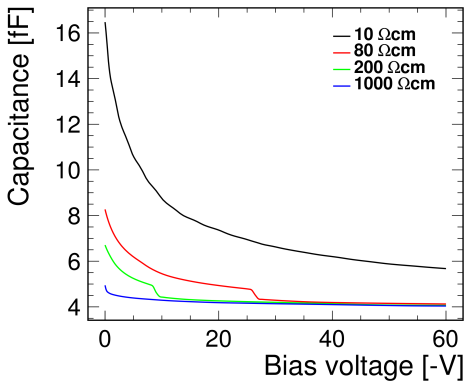
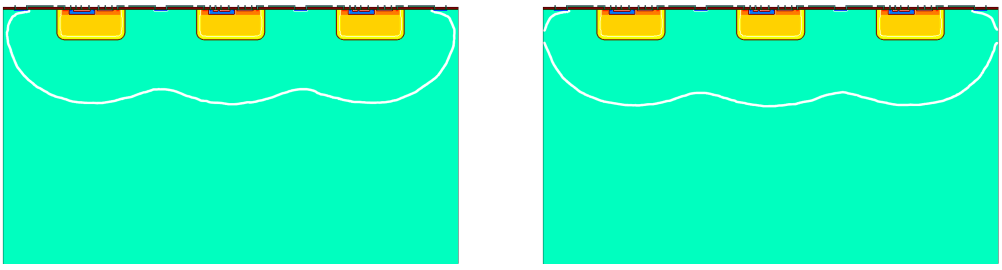


Figure B.1: The capacitance as a function of voltage for a three-pixel structure that is 81.5 µm wide.



(a) **(b)**

Figure B.2: The depletion region of the 80 Ω cm model for two voltages: (a) –24 V and (b) –26 V.

APPENDIX C

Bump bond quality plots

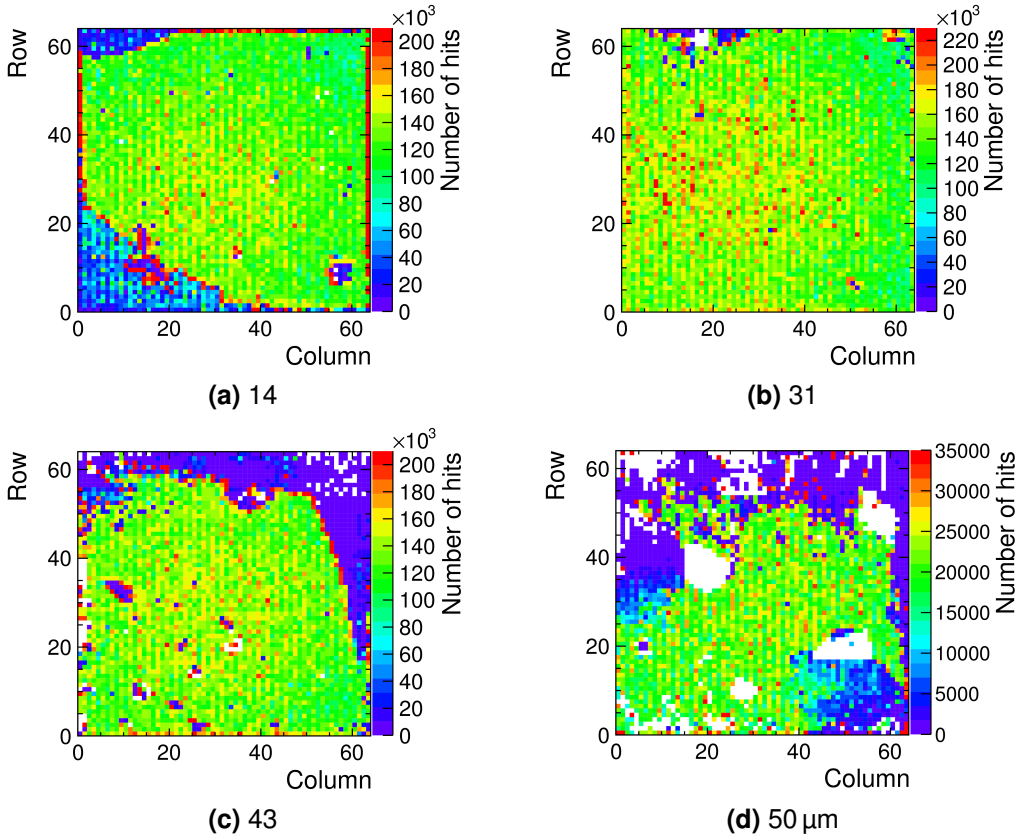
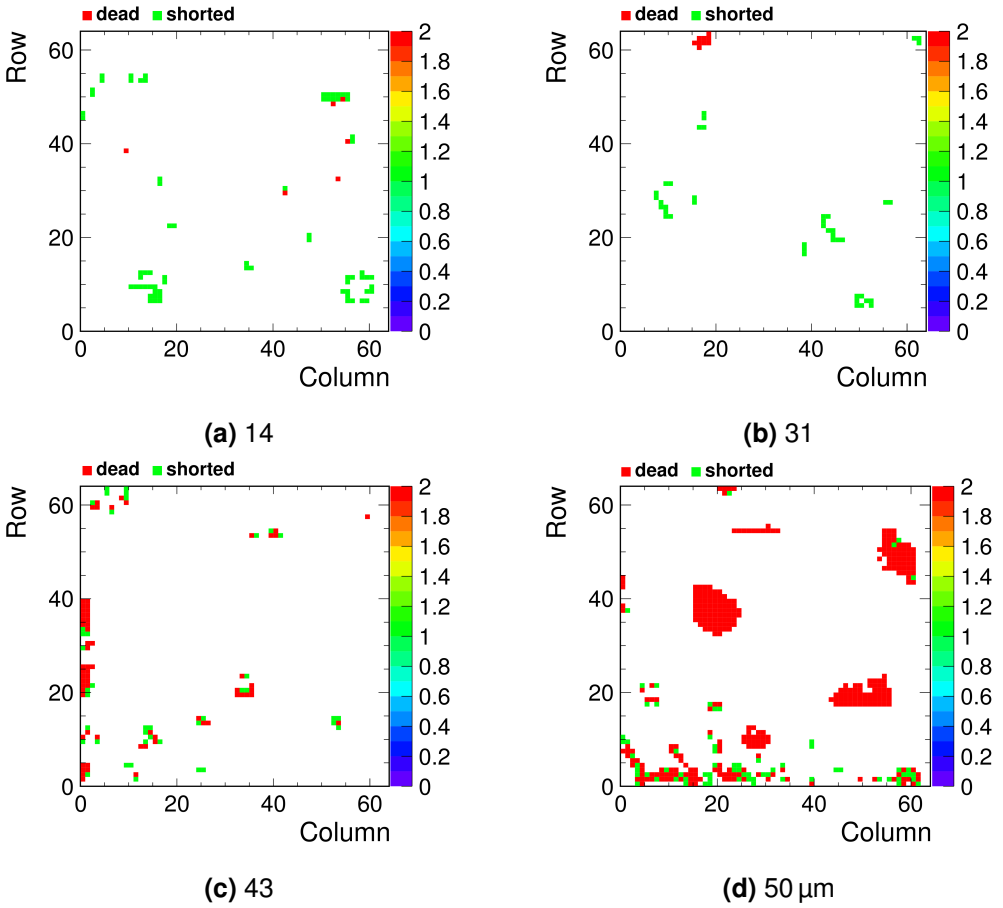


Figure C.1: The source hit maps of all planar assemblies.



(a) 14 **(b)** 31

(c) 43 **(d)** 50 µm

Figure C.2: The location of pixels connected to neighbours (green) and unresponsive to the test pulses (red) for all planar assemblies.

References

[1] G. Aad et al., ATLAS, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, DOI: 10.1016/j.physletb.2012.08.020.

[2] S. Chatrchyan et al., CMS, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, DOI: 10.1016/j.physletb.2012.08.021.

[3] F. Englert, R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321, DOI: 10.1103/PhysRevLett.13.321.

[4] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132, DOI: 10.1016/0031-9163(64)91136-9.

[5] M. Aicheler et al., *A Multi-TeV Linear Collider Based on CLIC Technology: CLIC Conceptual Design Report*, tech. rep. CERN-2012-007, Geneva, 2012.

[6] L. Linssen et al., *Physics and Detectors at CLIC: CLIC Conceptual Design Report* (2012).

[7] R. Pierret, *Advanced Semiconductor Fundamentals*, Modular series on solid state devices, Addison-Wesley Publishing Company, 1987, ISBN: 9780201053388.

[8] J. S. Blakemore, *Solid state physics; 2nd ed.* Cambridge Univ. Press, Cambridge, 1985.

[9] C. Kittel, *Introduction to Solid State Physics; 8th ed.* Wiley, Hoboken, NJ, 2005.

[10] C. Patrignani, P. D. Group, *Review of Particle Physics*, Chinese Physics C **40** (2016) 100001, URL: http://stacks.iop.org/1674-1137/40/i=10/a=100001.

[11] L. Landau, *On the energy loss of fast particles by ionization*, J. Phys.(USSR) **8** (1944) 201.

[12] L. Rossi et al., *Pixel detectors: from fundamentals to applications*, Particle Acceleration and Detection, Springer, Berlin, 2006.

[13] S. M. Sze, *Physics of semiconductor devices*, 2nd ed., Wiley, New York, NY, 1981.

[14] H.-G. Moser, *Silicon detector systems in high energy physics*, Prog. Part. Nucl. Phys. **63** (2009) 186, DOI: 10.1016/j.ppnp.2008.12.002.

[15] A. Dorokhov et al., *High resistivity CMOS pixel sensors and their application to the STAR PXL detector*, Nucl. Instrum. Methods Phys. Res. A: **650** (2011), International Workshop on Semiconductor Pixel Detectors for Particles and Imaging 2010 174, ISSN: 0168-9002, DOI: https://doi.org/10.1016/j.nima.2010.12.112.

[16] C. Cavicchioli et al., *Design and characterization of novel monolithic pixel sensors for the ALICE ITS upgrade*,
Nucl. Instrum. Methods Phys. Res. A: **765** (2014), HSTD-9 2013 - Proceedings of the 9th International Hiroshima Symposium on Development and Application of Semiconductor Tracking Detectors 177, ISSN: 0168-9002,
DOI: https://doi.org/10.1016/j.nima.2014.05.027.

[17] I. Perić, *Hybrid Pixel Particle Detector Without Bump Interconnection*,
Nucl. Sci. **56** (2009) 519.

[18] W. Snoeys, *CMOS monolithic active pixel sensors for high energy physics*,
Nucl. Instrum. Methods Phys. Res. A: **765** (2014), HSTD-9 2013 - Proceedings of the 9th International Hiroshima Symposium on Development and Application of Semiconductor Tracking Detectors 167, ISSN: 0168-9002,
DOI: https://doi.org/10.1016/j.nima.2014.07.017.

[19] G. Moortgat-Pick et al., *The Role of polarized positrons and electrons in revealing fundamental interactions at the linear collider*,
Phys. Rept. **460** (2008) 131, DOI: 10.1016/j.physrep.2007.12.003.

[20] H. Abramowicz et al.,
Higgs Physics at the CLIC Electron-Positron Linear Collider,
Eur. Phys. J. C **77** (2016), Comments: 40 pages, 28 figures 475. 40 p,
URL: http://cds.cern.ch/record/2210491.

[21] M. A. Thomson, *Model-independent measurement of the $e^+e^- \to HZ$ cross section at a future e^+e^- linear collider using hadronic Z decays*,
The European Physical Journal C **76** (2016) 72,
DOI: 10.1140/epjc/s10052-016-3911-5.

[22] K. Seidel et al., *Top quark mass measurements at and above threshold at CLIC*,
Eur. Phys. J. **C73** (2013) 2530, DOI: 10.1140/epjc/s10052-013-2530-7.

[23] M. J. Boland et al., CLIC and CLICdp collaborations,
Updated baseline for a staged Compact Linear Collider,
tech. rep. CERN-2016-004, Geneva, 2016.

[24] P. Lebrun et al., *The CLIC Programme: Towards a Staged e^+e^- Linear Collider Exploring the Terascale: CLIC Conceptual Design Report*,
tech. rep. CERN-2012-005, Geneva, 2012.

[25] *CLICdp*,
URL: http://clicdp.web.cern.ch/content/clic-accelerator
(visited on 03/01/2018).

[26] P. Chen, *Beamstrahlung and the QED, QCD backgrounds in linear colliders* (1992) 12 p, URL: https://cds.cern.ch/record/242895.

[27] P. Chen, T. L. Barklow, M. E. Peskin, *Hadron production in gamma gamma collisions as a background for e+ e- linear colliders*,
Phys. Rev. **D49** (1994) 3209, DOI: 10.1103/PhysRevD.49.3209.

[28] N. Alipour Tehrani et al., *CLICdet: The post-CDR CLIC detector model*,
CLICdp-Note-2017-001, 2017.

[29] A. M. Nurnberg, D. Dannheim, *Requirements for the CLIC tracker readout* (2017), URL: https://cds.cern.ch/record/2261066.

[30] M. A. Thomson, *Particle flow calorimetry and the PandoraPFA algorithm*, Nuclear Instruments and Methods in Physics Research A **611** (2009) 25.

[31] T. C. collaboration, J. Repond et al.,
Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider,
Journal of Instrumentation **3** (2008) P08001.

[32] C. Adloff et al., CALICE, *Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype*, JINST **5** (2010) P05004.

[33] H. Abramowicz et al., *Forward instrumentation for ILC detectors*,
Journal of Instrumentation **5** (2010) P12002.

[34] T. Price et al.,
Full simulation study of the top Yukawa coupling at the ILC at $\sqrt{s}$=1 TeV,
75 (2014).

[35] S. Dawson et al., *Working Group Report: Higgs Boson*, Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013, 2013, URL: https://inspirehep.net/record/1262795/files/arXiv:1310.8361.pdf.

[36] R. S. Gupta, H. Rzehak, J. D. Wells,
How well do we need to measure the Higgs boson mass and self-coupling?,
Phys. Rev. D **88** (5 2013) 055024, DOI: 10.1103/PhysRevD.88.055024.

[37] R. S. Gupta, H. Rzehak, J. D. Wells,
How well do we need to measure Higgs boson couplings?,
Phys. Rev. D **86** (9 2012) 095001, DOI: 10.1103/PhysRevD.86.095001.

[38] M. Vos et al., *Top physics at high-energy lepton colliders* (2016),
arXiv: 1604.08122 [hep-ex].

[39] A. Zarnecki, N. Van Der Kolk, CLICdp Collaboration,
Limits on top FCNC decay $t\to cH$ and $t\to cP\gamma$ from CLIC at 380 GeV,
tech. rep. CLICdp-Conf-2018-001,
9 pages, 6 figures. Talk presented at the International Workshop on Future Linear Colliders (LCWS2017), Strasbourg, France, 23-27 October 2017. C17-10-23.2. Presented on behalf of the CLICdp Collaboration,
Geneva: CERN, 2018, URL: http://cds.cern.ch/record/2300302.

[40] V. Khachatryan et al., CMS, *Measurement of the top quark mass using proton-proton data at $\sqrt{s}$ = 7 and 8 TeV*, Phys. Rev. **D93** (2016) 072004,
DOI: 10.1103/PhysRevD.93.072004, arXiv: 1509.04044 [hep-ex].

[41] C. Collaboration, CMS, *Projected improvement of the accuracy of top-quark mass measurements at the upgraded LHC* (2013).

[42] G. Cortiana, *Top-quark mass measurements: review and perspectives*,
Rev. Phys. **1** (2016) 60, DOI: 10.1016/j.revip.2016.04.001.

[43] M. S. Amjad et al.,
A precise characterisation of the top quark electro-weak vertices at the ILC,
Eur. Phys. J. **C75** (2015) 512, DOI: 10.1140/epjc/s10052-015-3746-5,
arXiv: 1505.06020 [hep-ex].

[44] N. Van Der Kolk, *Toward Precision Top Quark Measurements in e^+e^- Collisions at Linear Colliders*, tech. rep. CLICdp-Conf-2017-016, Geneva: CERN, 2017, URL: http://cds.cern.ch/record/2287576.

[45] ATLAS Collaboration, *Expected sensitivity of ATLAS to FCNC top quark decays t→Zu and t→Hq at the High Luminosity LHC*, tech. rep. ATL-PHYS-PUB-2016-019, Geneva: CERN, 2016, URL: https://cds.cern.ch/record/2209126.

[46] F. Duarte Ramos, W. Klempt, F.-X. Nuiry, *Experimental tests on the air cooling of the CLIC vertex detector* (2016), URL: http://cds.cern.ch/record/2138963.

[47] G. Blanchot, C. A. Fuentes Rojas, *Power pulsing scheme for analog and digital electronics of the vertex detectors at CLIC* (2015), URL: http://cds.cern.ch/record/2062429.

[48] D. Dannheim, A. Sailer, *Beam-Induced Backgrounds in the CLIC Detectors* (2012), URL: https://cds.cern.ch/record/1443516.

[49] F. Duarte Ramos, H. Gerwig, M. Villarejo Bermudez, CERN Linear Collider Detector collaboration, *CLIC inner detectors cooling simulations* (2014), URL: http://cds.cern.ch/record/1572989.

[50] P. Valerio, R. Ballabriga, M. Campbell, CERN Linear Collider Detector collaboration, *Design of the 65 nm CLICpix demonstrator chip* (2012), URL: https://cds.cern.ch/record/1507691.

[51] X. Llopart et al., *Timepix, a 65k programmable pixel readout chip for arrival time, energy and/or photon counting measurements*, Nucl. Instrum. Meth. **A581** (2007) 485.

[52] N. A. Tehrani et al., *Capacitively coupled hybrid pixel assemblies for the CLIC vertex detector*, Nucl. Instrum. Methods Phys. Res. A: **823** (2016) 1, ISSN: 0168-9002, DOI: https://doi.org/10.1016/j.nima.2016.03.072.

[53] P. Valerio et al., *A prototype hybrid pixel detector ASIC for the CLIC experiment* (2013), URL: https://cds.cern.ch/record/1623863.

[54] *Micron*, URL: https://www.micron.com/ (visited on 02/01/2018).

[55] *Advacam*, URL: http://advacam.com/ (visited on 02/01/2018).

[56] T. Poikela et al., *Timepix3: a 65K channel hybrid pixel readout chip with simultaneous ToA/ToT and sparse readout*, Journal of Instrumentation **9** (2014) C05013, URL: http://stacks.iop.org/1748-0221/9/i=05/a=C05013.

[57] N. Alipour Tehrani, On behalf of the CLICdp collaboration, *Recent results with HV-CMOS and planar sensors for the CLIC vertex detector*, tech. rep. CLICdp-Conf-2016-005, Geneva: CERN, 2016, URL: http://cds.cern.ch/record/2143800.

[58] A. Tomada et al., *Flip Chip Assembly of Thin Substrates, Fine Bump Pitch, and Small Prototype Die* (2014), SLAC-PUB-16168.

[59] M. Dinardo, *The pixel detector for the CMS phase-II upgrade*, Journal of Instrumentation **10** (2015) C04019, URL: http://stacks.iop.org/1748-0221/10/i=04/a=C04019.

[60] I. Perić, C. Kreidl, P. Fischer, *Hybrid pixel detector based on capacitive chip to chip signal-transmission*, Nucl. Instrum. Meth. **A617** (2010) 576.

[61] I. Kremastiotis et al., *Design and standalone characterisation of a capacitively coupled HV-CMOS sensor chip for the CLIC vertex detector*, Journal of Instrumentation **12** (2017) P09012, URL: http://stacks.iop.org/1748-0221/12/i=09/a=P09012.

[62] I. Perić et al., *High-voltage pixel detectors in commercial CMOS technologies for ATLAS, CLIC and Mu3e experiments*, Nucl. Instrum. Meth. **A731** (2013) 131.

[63] *ACCµRA 100 flip-chip bonder supplied by Setna.* URL: http://set-na.com/products/acc100/ (visited on 25/09/2017).

[64] *Technical datasheet Uhu plus endfest 300 epoxy resin.* URL: https://edms.cern.ch/document/1517626/ (visited on 25/09/2017).

[65] *CERN engineering department, mechanical and materials engineering, materials, metrology & NDT section (CERN-EN-MME-MM).* URL: http://en.web.cern.ch/en-mme-group (visited on 25/09/2017).

[66] C. Maiti, *Introducing Technology Computer-Aided Design (TCAD): Fundamentals, Simulations, and Applications*, Pan Stanford Publishing, 2017, ISBN: 9789814745529.

[67] L. W. Nagel, D. Pederson, *SPICE (Simulation Program with Integrated Circuit Emphasis)*, tech. rep. UCB/ERL M382, EECS Department, University of California, Berkeley, 1973.

[68] *Synopsys TCAD*, URL: https://www.synopsys.com/silicon/tcad.html (visited on 31/05/2017).

[69] *Sentaurus Structure Editor User Guide*, Version M-2016.12, Synopsys, 2016, URL: https://solvnet.synopsys.com/dow_retrieve/M-2017.06/tcad/pdf/sense_ug.pdf?otSearchResultSrc=advSearch&otSearchResultNumber=5&otPageNum=1.

[70] *Sentaurus Device User Guide*, Version M-2016.12, Synopsys, 2016, URL: https://solvnet.synopsys.com/dow_retrieve/M-2017.06/tcad/pdf/sdevice_ug.pdf?otSearchResultSrc=advSearch&otSearchResultNumber=1&otPageNum=1.

[71] *Sentaurus TCAD: Industry-Standard Process and Device Simulators*, URL: https://www.synopsys.com/content/dam/synopsys/silicon/datasheets/sentaurus_ds.pdf (visited on 15/06/2017).

[72] *Stanford TCAD, Stanford University*, URL: http://www-tcad.stanford.edu (visited on 12/06/2017).

[73] S. Li, Y. Fu, *3D TCAD Simulation for Semiconductor Processes, Devices and Optoelectronics; 1st ed.* Springer, New York, 2012.

[74] W. Schoenmaker et al.,
Renormalization group meshes and the discretization of TCAD equations,
IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **21** (2002) 1425, ISSN: 0278-0070,
DOI: 10.1109/TCAD.2002.804383.

[75] Y. Fu, Z. Li, W. T. Ng, *Integrated power devices and TCAD simulation*,
Devices, circuits, and systems, CRC Press, London, 2014.

[76] *Ligament User Guide*, Version M-2016.12, Synopsys, 2016,
URL: https://solvnet.synopsys.com/dow_retrieve/M-2017.06/tcad/pdf/ligament_ug.pdf?otSearchResultSrc=advSearch&otSearchResultNumber=1&otPageNum=1.

[77] I. Kremastiotis, *Update on extracted simulations and C3PD design status*,
Presentation at the CLIC WG vertex meeting, CERN, Geneva, Switzerland, 2015, URL:
https://indico.cern.ch/event/452100/contributions/1960748/attachments/1199064/1744036/ExtractedDesignStatus.pdf.

[78] M. F. García et al., *High-resolution three-dimensional imaging of a depleted CMOS sensor using an edge Transient Current Technique based on the Two Photon Absorption process (TPA-eTCT)*, Nucl. Instrum. Methods Phys. Res. A: **845** (2017), Proceedings of the Vienna Conference on Instrumentation 2016 69, ISSN: 0168-9002,
DOI: https://doi.org/10.1016/j.nima.2016.05.070.

[79] S. Kulis, *uASIC documentation*, available at http://uasic.web.cern.ch/uasic/, 2014.

[80] P. Valerio et al., on behalf of The CLIC Detector and Physics Study (CLICdp),
A prototype hybrid pixel detector ASIC for the CLIC experiment,
JINST **9** (2013) C01012. 10,
URL: https://cds.cern.ch/record/1635171.

[81] M. D. Gaspari et al., *Design of the analog front-end for the Timepix3 and Smallpix hybrid pixel detectors in 130 nm CMOS technology*,
Journal of Instrumentation **9** (2014) C01037,
URL: http://stacks.iop.org/1748-0221/9/i=01/a=C01037.

[82] J. Jakubek et al., *Pixel detectors for imaging with heavy charged particles*,
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **591** (2008) 155,
ISSN: 0168-9002,
DOI: https://doi.org/10.1016/j.nima.2008.03.091.

[83] S. Green, M. Thomson, J. Marshall, *Calorimetry at a Future Linear Collider*,
Presented 11 Oct 2017, 2017,
URL: https://cds.cern.ch/record/2293158.

[84] I. Rubinskiy, H. Perrey,
An EUDET/AIDA Pixel Beam Telescope for Detector Development,
PoS **TIPP2014** (2014) 122.

[85] C. Hu-Guo et al., *First reticule size MAPS with digital output and integrated zero suppression for the EUDET-JRA1 beam telescope*,
Nucl. Instrum. Meth. **A623** (2010) 480, DOI: 10.1016/j.nima.2010.03.043.

[86] M. Karagounis, *Development of the ATLAS FE-I4 pixel readout IC for b-layer Upgrade and Super-LHC* (2008),
URL: https://cds.cern.ch/record/1158505.

[87] S. Straulino et al., *Spatial Resolution of Double-Sided Silicon Microstrip Detectors for the PAMELA Apparatus*,
Nucl. Instrum. Methods Phys. Res., A **556** (2005) 100,
URL: http://cds.cern.ch/record/892911.

[88] R. Turchetta, *Spatial resolution of silicon microstrip detectors*,
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **335** (1993) 44,
ISSN: 0168-9002,
DOI: http://dx.doi.org/10.1016/0168-9002(93)90255-G.

[89] K. Akiba et al., *Charged Particle Tracking with the Timepix ASIC*,
Nucl. Instrum. Meth. **A661** (2012) 31, DOI: 10.1016/j.nima.2011.09.021.

[90] *Cadence Virtuoso Analog Design Environment*,
URL: https://www.cadence.com/content/cadence-www/global/en_US/home/tools/custom-ic-analog-rf-design/circuit-design/virtuoso-analog-design-environment.html (visited on 04/10/2017).

[91] M. Vicente Barreto Pinto, *Finite-element simulations of coupling capacitances in capacitively coupled pixel detectors*, CLICdp-Note-2017-003, 2017.

[92] S. Reggiani et al., *Experimental extraction of the electron impact-ionization coefficient at large operating temperatures*,
IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.
P. 407, DOI: 10.1109/IEDM.2004.1419171.

[93] H. Jansen et al., *Performance of the EUDET-type beam telescopes*,
EPJ Tech. Instrum. **3** (2016) 7, DOI: 10.1140/epjti/s40485-016-0033-2.


LUX

UNIVERSITY OF
LIVERPOOL

SIMULATION AND EVALUATION OF HV-CMOS PIXEL SENSORS FOR THE CLIC VERTEX DETECTOR

Thesis submitted in accordance with the requirements
of the University of Liverpool for the degree of
Doctor of Philosophy
by

Matthew Buckland*†

Supervisors:
Prof. Joost Vossebeld*
Dr. Dominik Dannheim†

*University of Liverpool, United Kingdom
†CERN, Switzerland

June 2018